DIMETHYLALLYLTYROSINE SYNTHASE AND DIMETHYLALLYLTRYPTOPHAN SYNTHASE: PROMISCUOUS AROMATIC PRENYLTRANSFERASES

by

Jeffrey Daniel Rudolf

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

May 2013

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Jeffrey Daniel Rudolf**

has been approved by the following supervisory committee members:

C. Dale Poulter	, Chair	**12-7-2012** Date Approved
Cynthia J. Burrows	, Member	**12-7-2012** Date Approved
Kenneth J. Woycechowsky	, Member	**12-7-2012** Date Approved
John C. Conboy	, Member	**12-7-2012** Date Approved
Christopher P. Hill	, Member	**12-7-2012** Date Approved

and by **Henry S. White**, Chair of the Department of **Chemistry**

and by Donna M. White, Interim Dean of The Graduate School.

ABSTRACT

The aromatic prenyltransferase SirD in *Leptosphaeria maculans* shares sequence homology with fungal indole prenyltransferases and is hypothesized to catalyze the normal *O*-prenylation of tyrosine in the first committed step in the biosynthesis of the phytotoxin sirodesmin PL. We cloned, expressed, and biochemically characterized SirD and discovered that it was a dimethylallyltyrosine synthase. We synthesized tyrosine and tryptophan derivatives and tested SirD for aromatic substrate promiscuity. NMR studies revealed that SirD catalyzes normal *O*-, *N*-, and *S*-prenylations on tyrosine derivatives and *N*- and *C*-prenylations on tryptophan derivatives with normal prenylation at C6 and C7 and reverse prenylation at N1 and C7 of the indole ring, depending on the substituent.

The aromatic prenyltransferase dimethylallyltryptophan synthase in *Claviceps purpurea* catalyzes the normal prenylation of tryptophan at C4 of the indole nucleus in the first committed step of ergot alkaloid biosynthesis. 4-Methyltryptophan is a competitive inhibitor of the enzyme that has been used in kinetic studies. We found that 4-methyltryptophan is an alternate substrate, which gave four products. Similarly, 4-methoxytryptophan and 4-aminotryptophan are alternate substrates. NMR studies revealed that dimethylallyltryptophan synthase catalyzed normal prenylation at N1, C3, C5, and C7 and reverse prenylation at C3 of the indole ring, depending on the substituent. Synthesized aromatic amino acid analogues were also tested with dimethylallyl-tryptophan synthase.

The ability of dimethylallyltyrosine synthase and dimethylallyltryptophan synthase to prenylate six different sites on the indole nucleus, with normal and reverse prenylation at two of the sites, is consistent with a dissociative electrophilic alkylation of the indole ring where orientation of the substrates within the active site and substituent electronic effects determine the position and type of prenylation.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS xi

ACKNOWLEDGMENTS xiii

CHAPTER

1. INTRODUCTION 1

Isoprenoids 2
Isoprenoid Biosynthesis 2
Secondary Metabolites 5
Prenyltransferases 9
Dimethylallyltyrosine Synthase 14
Dimethylallyltryptophan Synthase 16
Dissertation Overview 20

2. SYNTHESIS OF ENZYMATIC SUBSTRATES AND AROMATIC AMINO ACID ANALOGUES 22

Introduction 23
Experimental Procedures 24
Results and Discussion 43

3. SIRD: BIOCHEMICAL CHARACTERIZATION AND AROMATIC PROMISCUITY 50

Introduction 51
Experimental Procedures 53
Results and Discussion 68

4. 4-DMATS: MULTISITE PRENYLATION OF 4-SUBSTITUTED TRYPTOPHANS 109

Introduction 110
Experimental Procedures 112
Results and Discussion 119
Conclusion 142

APPENDICES

A. HRMS DATA 145

B. NMR SPECTRA 147

REFERENCES 271

LIST OF TABLES

Table Page

3.1. Crystallization conditions resulting in SirD crystal formation 74

3.2. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from L-tyrosine and 4-mercapto-L-phenylalanine 80

3.3. Kinetic parameters for SirD substrates at 30 °C 85

3.4. Aromatic substrate analogues tested for turnover with SirD and corresponding product ratios 90

3.5. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from L-tryptophan and 4-methyl-L-tryptophan 95

3.6. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from 4-methoxy and 7-methyl-L-tryptophan 102

4.1. Aromatic substrate analogues tested for turnover with 4-DMATS and the corresponding product ratios 123

4.2. Kinetic parameters for 4-DMATS substrates at 30 °C 124

4.3. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from 4-methyl and 4-methoxy-L-tryptophan 127

4.4. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from 4-amino and 2-oxo-L-tryptophan 134

S.1. HRMS data of synthesized compounds 146

LIST OF FIGURES

Figure Page

1.1. Chemical structures of representative examaples of ETPs and prenylated alkaloids. 7

1.2. Chemical structures of representative prenyltransferase products highlighting the different types of prenyl acceptors. 10

1.3. Cartoon representation of the 4-DMATS (FgaPT2) crystal structure 19

1.4. 4-DMATS proposed mechanisms. 21

2.1. Amino acid analogues synthesized 49

3.1. The engineered sirD gene constructs synthesized by Genscript. 54

3.2. *L. maculans* sirD engineered gene and encoded amino acid sequence. 69

3.3. *S. diversum* sirD engineered gene and encoded amino acid sequence. 70

3.4. SDS-PAGE of pLmSirD expression in BL21 (DE3). 71

3.5. SDS-PAGE of pSdSirD expression in BL21 (DE3). 73

3.6. Diffraction pattern obtained from the ~8 Å SirD JENA D3 crystal 74

3.7. A sequence alignment of *L. maculans* SirD and *A. fumigatus* FgaPT2 (4-DMATS) containing the predicted secondary structure of SirD and the secondary structure determined by FgaPT2's crystal structure 76

3.8. Homology structure of SirD. 77

3.9. Active site of SirD 77

3.10. Radioautography TLC of the assay for the determination SirD's natural substrates. 79

3.11. Thermal stability of SirD 82

3.12. Michaelis-Menten kinetic plots for select SirD substrates 84

3.13. Radioautography TLCs of substrate promiscuity assays for SirD 86

3.14. Radioautography TLCs of substrate promiscuity assays for SirD 88

3.15. HPLC chromatogram of 4-mercapto-L-phenylalanine and DMAPP incubation with SirD 92

3.16. HPLC chromatogram of 4-methyl-L-tryptophan and DMAPP incubation with SirD 93

3.17. HPLC chromatogram of L-tryptophan and DMAPP incubation with SirD 97

3.18. Timecourse of L-tryptophan or 4-methyl-L-tryptophan incubations with DMAPP and SirD 98

3.19. HPLC chromatogram of prenylated 4-methyl-L-tryptophan and DMAPP incubation with SirD 99

3.20. HPLC chromatogram of 7-methyl-L-tryptophan and DMAPP incubation with SirD 101

3.21. HPLC chromatogram of 4-methoxy-L-tryptophan and DMAPP incubation with SirD 104

3.22. Tryptophan prenylation sites and orientations catalyzed by SirD. 106

3.23. HPLC chromatogram of 4-vinyl-L-tryptophan and DMAPP incubation with SirD 107

3.24. HPLC chromatogram of DIT and DMAPP incubation with SirD 108

4.1. SDS-PAGE of pHDMAT/pGroESL co-expression in BL21 (DE3) 120

4.2. Radioautography TLCs of substrate promiscuity assays for 4-DMATS 121

4.3. Michaelis-Menten kinetic plots for select 4-DMATS substrates 124

4.4. HPLC chromatogram of 4-methyl-L-tryptophan and DMAPP incubation with 4-DMATS 126

4.5. HPLC chromatograms of prenylated 4-methyltryptophan product co-injection 130

4.6. HPLC chromatogram of 4-methoxy-L-tryptophan and DMAPP incubation with 4-DMATS 131

4.7. HPLC chromatogram of 4-amino-L-tryptophan and DMAPP incubation with 4-DMATS 133

4.8. Prenylation sites and orientations catalyzed by 4-DMATS. 136

4.9. HPLC chromatogram of 4-hydroxy-L-tryptophan and DMAPP incubation with 4-DMATS 137

4.10. HPLC chromatogram of dihydroiso-L-tryptophan and DMAPP incubation with 4-DMATS 139

4.11. HPLC chromatograms of prenylated 2-oxo-L-tryptophan incubations with 4-DMATS 140

4.12. HPLC chromatogram of 2-oxo-L-tryptophan and DMAPP incubation with 4-DMATS 141

LIST OF ABBREVIATIONS

2D	two-dimensional
ACN	acetonitrile
BLAST	basic local alignment search tool
CDP-ME	4-diphosphocytidyl-2-*C*-methyl-D-erythritol
COSY	correlation spectroscopy
CTP	cytidine triphosphate
Da	Dalton
DMAPP	dimethylallyl diphosphate
DMASPP	dimethylallyl *S*-thiolodiphosphate
DMAT	dimethylallyltryptophan
DMATS	dimethylallyltryptophan synthase
DMSO	dimethyl sulfoxide
ETP	epipolythiodioxopiperazine
Fmoc	fluorenylmethyloxycarbonyl
FPP	farnesyl diphosphate
GPP	geranyl diphosphate
GGPP	geranylgeranyl diphosphate
HHPI	hexahydropyrroloindole
HMBC	heteronuclear multiple-bond correlation spectroscopy

HMQC	heteronuclear multiple-quantum correlation spectroscopy
HPLC	high performace liquid chromatography
HRMS	high-resolution mass spectrometry
HSQC	heteronuclear single-quantum correlation spectroscopy
IDI	isopentenyl diphosphate isomerase
IPP	isopentenyl diphosphate
IPTG	isopropyl-β-D-thiogalactopyranoside
MEP	2-*C*-methyl-D-erythritol 4-phosphate
MVA	mevalonate
NMR	nuclear magnetic resonance
PLP	pyridoxal-5'-phosphate
ROESY	rotating frame nuclear Overhauser effect spectroscopy
RP	reverse phase
t-Boc	*tert*-butoxycarbonyl
TFA	trifluoroacetic acid
TLC	thin layer chromatography
TOCSY	total correlation spectroscopy

ACKNOWLEDGMENTS

First and foremost, thanks to Dr. C. Dale Poulter for his mentorship, advice, support, patience and trust throughout my time in his lab. Thanks to all the past and present members of the Poulter group for their helpful discussions, enlightened ideas, and friendship. Special thanks to my extremely patient and loving wife, Becky, who tolerated my extended hours, weekdays and weekends, in the lab.

This work was supported by NIH grant GM 21328.

CHAPTER 1

INTRODUCTION

Isoprenoids

Isoprenoids, commonly called terpenoids, constitute the largest known family of natural products. According to The Dictionary of Natural Products, over 58,000 compounds of this class are reported.[1] The structural diversity necessary to create this vast array of distinct compounds originates from the use of the five-carbon building block isoprene. It is a fascinating example of nature's creativity; using a simple five-carbon chain to build larger, structurally more complex carbon backbones, which are further tailored by the organism.

Synthesized and exploited by all three domains of life (Eukaryota, Bacteria, and Archaea), isoprenoids serve as an assortment of functional tools.[2-5] Functional roles are classified into two categories – essential roles (primary metabolism) and nonessential roles (secondary metabolism). Examples of essential roles include: membrane integrity (cholesterol, diphytanyldiglycerol tetraether lipids in archaea[2]), hormones (testosterone, estrogen, gibberellins), electron carriers (ubiquinone, heme A), cellular localization and lipid anchoring (protein prenylation), and photosynthetic pigments (carotenoids). Examples of secondary metabolism roles include: attractants for pollinators or seed dispersers (flavonoids),[6] defense mechanisms (ergot alkaloids),[6,7] and virulence (sirodesmin PL).[8] Isoprenoids are also useful as renewable resources and commercial commodities, including polymers, flavorings and fragrances, adhesives and solvents, as well as agricultural chemicals (pyrethrins) and pharmaceuticals (taxol, artemisinin).[3]

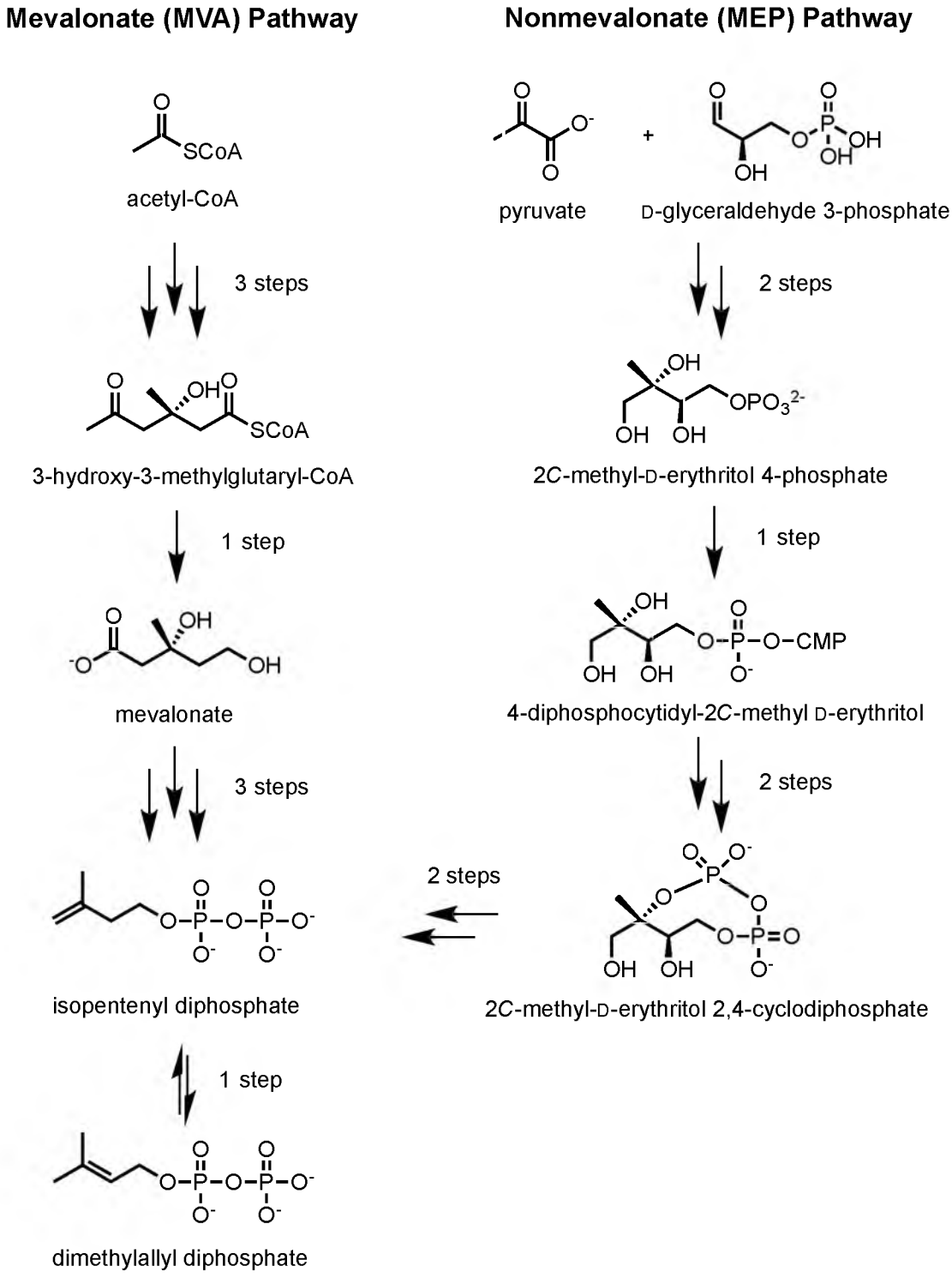
Isoprenoid Biosynthesis

In Nature, isoprenoids are derived from the five-carbon building block isoprene, but organisms do not use isoprene directly. Instead, two isoprene derivatives, isopentenyl

diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), are utilized as activated forms of the five-carbon building block. Organisms use one of two pathways, and sometimes both, to synthesize the universal precursors IPP and DMAPP (Scheme 1.1).[9] The mevalonate (MVA) pathway is present in higher eukaryotes, archaea, and some bacteria;[10] the nonmevalonate pathway, or 2*C*-methyl-D-erythritol-4-phosphate (MEP) pathway, is an alternative pathway used in bacteria, plant chloroplasts, and algae.[11,12]

The MVA pathway condenses three molecules of acetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) before reduction to mevalonate. Two subsequent phosphorylations and a decarboxylation yields IPP.[10] In the final step, isopentenyl diphosphate isomerase (IDI) converts IPP to DMAPP.[13] In the MEP pathway, pyruvate and glyceraldehyde 3-phosphate are condensed and reduced to form 2*C*-methyl-D-erythritol 4-phosphate. MEP then reacts with CTP to form CDP-ME, which is phosphorylated and cyclized to form a methylerythritol cyclodiphosphate. Subsequent reduction and water elimination yields IPP and some DMAPP. To control the titer of IPP and DMAPP, IDI is also utilized in the MEP pathway.[11] The presence of DMAPP, as well as other allylic isoprenoids, is essential for isoprenoid chemistry as the allylic double bond stabilizes the positive charge formed by diphosphate dissociation.

To create structural diversity from IPP and DMAPP, the five-carbon units are connected via prenyltransferase reactions to produce longer, linear isoprenoids. The principal chain elongation reaction is a 1'-4 condensation between IPP (prenyl acceptor) and DMAPP (prenyl donor) to form the C_{10} isoprenoid geranyl diphosphate (GPP).[13] Successive additions of IPP form farnesyl diphosphate (FPP, C_{15}), geranylgeranyl diphosphate (GGPP, C_{20}), and so on.[14] These longer isoprenoids are then used as branch


Mevalonate (MVA) Pathway
Nonmevalonate (MEP) Pathway
acetyl-CoA
pyruvate
D-glyceraldehyde 3-phosphate
3 steps
2 steps
3-hydroxy-3-methylglutaryl-CoA
2C-methyl-D-erythritol 4-phosphate
1 step
1 step
mevalonate
4-diphosphocytidyl-2C-methyl D-erythritol
3 steps
2 steps
2 steps
isopentenyl diphosphate
2C-methyl-D-erythritol 2,4-cyclodiphosphate
1 step
dimethylallyl diphosphate

Scheme 1.1

points for the diverse array of biosynthetic pathways found in all domains of life (Scheme 1.2).

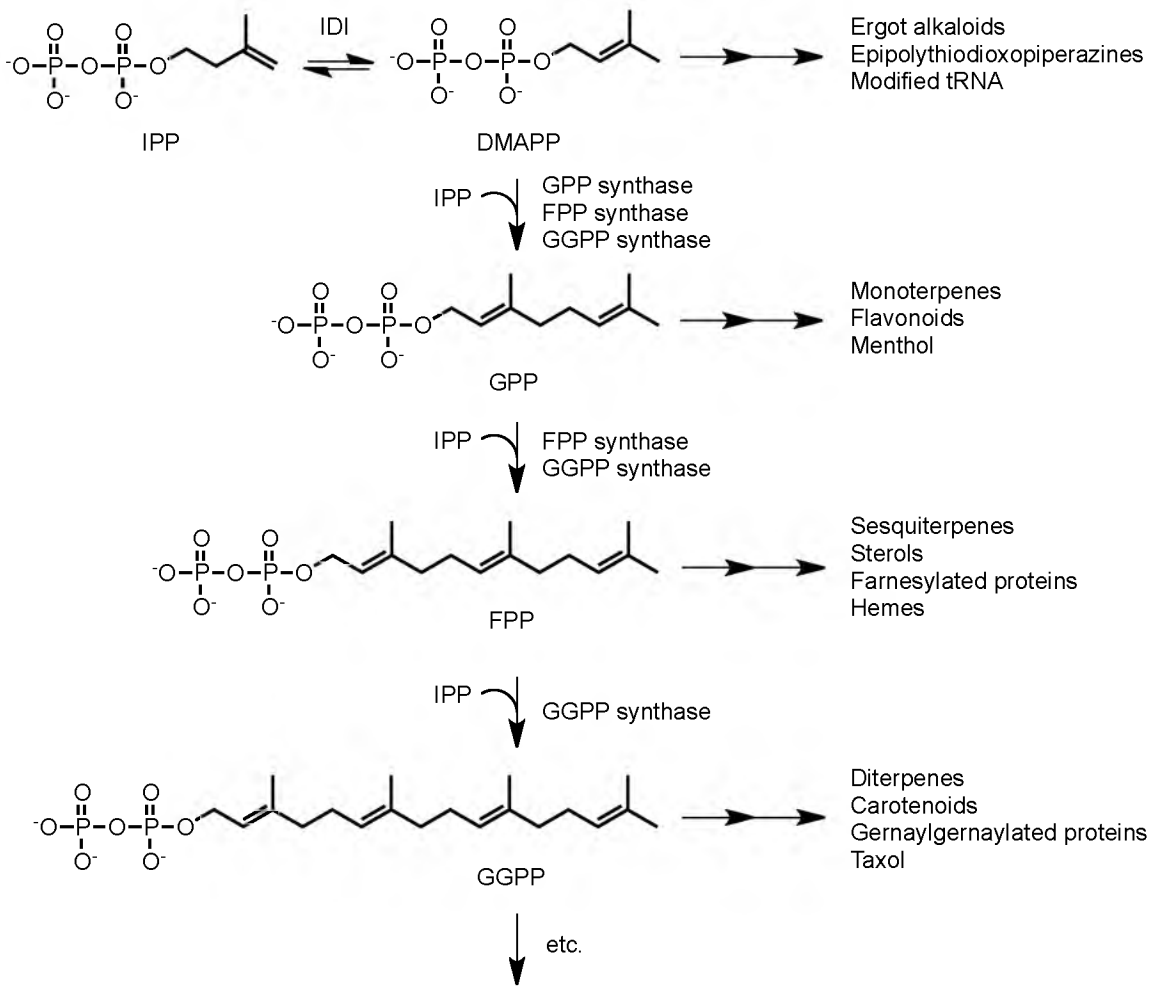
Secondary Metabolites

Isoprenoids are categorized into essential roles (primary metabolism) or non-essential roles (secondary metabolism). Primary metabolism provides metabolites necessary for an organism's survival. Secondary metabolites, while typically produced from primary metabolism intermediates, are compounds superfluous for an organism's survival. They may, however, provide a selective advantage to the producing organism to facilitate survival in an ecological niche.[15] Secondary metabolism is common in bacteria, plants, and fungi. Examples of secondary metabolite classes include polyketides, nonribosomal peptides (dioxopiperazines), and terpenes (ergot alkaloids).[16]

Epipolythiodioxopiperazines (ETPs) are fungal secondary metabolites characterized by a sulfur-bridged dioxopiperazine ring formed from the condensation of two amino acids (Figure 1.1).[17] Although this core moiety is found to carry various modifications, the molecular mechanism of toxicity originates from the sulfur atoms in the bridged ring. Two mechanisms are hypothesized to confer this toxicity – ETP-protein cross-linking through cysteine residues causing enzyme inactivation and reactive oxygen species production through redox cycling.[18] Although the exact function of ETPs is unknown, their toxicity suggests they are chemical agents used for defense or virulence. This toxicity also makes ETPs attractive pharmaceutical options.

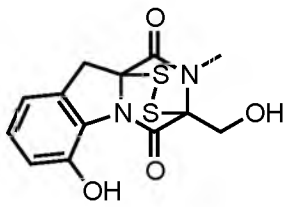
Gliotoxin, the first reported ETP,[19] is produced from many different fungal species including *Gliocladium fimbriatum* and *Aspergillus fumigatus*. It causes both apoptotic and necrotic cell death complemented with the generation of reactive oxygen

Isoprenoid Biosynthesis

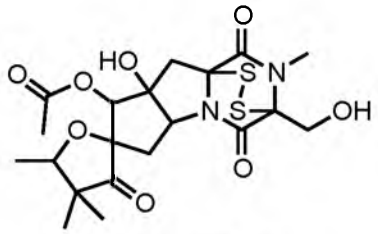


Scheme 1.2

ETPs

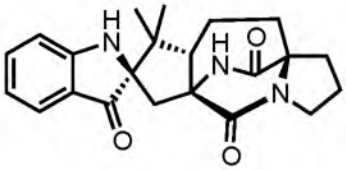


gliotoxin

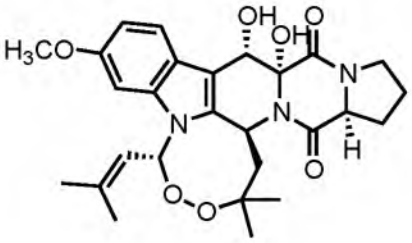


sirodesmin PL

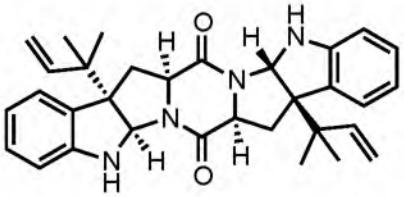
Prenyated Indole Alkaloids



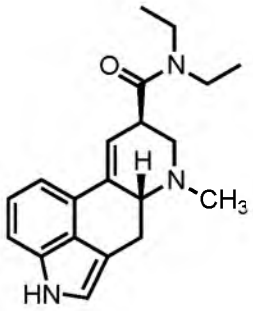
brevianamide A



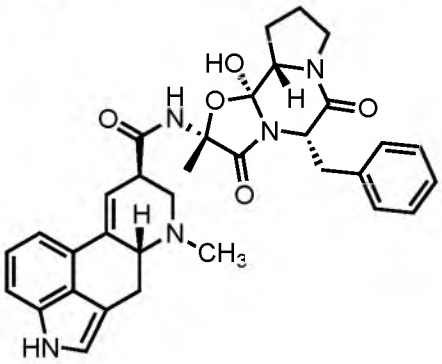
verruculogen



epiamauromine



lysergic acid diethylamide



ergotamine

Figure 1.1. Chemical structures of representative examaples of ETPs and prenylated alkaloids.

species and mitochondrial membrane disruption.[17] Gliotoxin was also shown to inhibit the protein farnesyltransferase and protein geranylgeranyltransferase enzymes, introducing ETPs as potential anti-cancer therapeutics.[20] *Leptosphaeria maculans*, the causal agent of blackleg disease in canola,[21] produces sirodesmin PL, a phytotoxin that causes host nonspecific necrosis.[8,21] Interestingly, sirodesmin PL is an ETP-terpene hybrid compound. *Sirodesmin diversum* produces chiral isomers of sirodesmin PL.[17] Other ETPs include sporidesmin A, the *O*-prenylated dithiosilvatin, and bis-ETPs such as leptosin.[18]

Prenylated indole alkaloids constitute a large family of chemically diverse biologically active secondary metabolites (Figure 1.1).[22-24] Biosynthetically, the carbon skeletons of these compounds are constructed from prenyl diphosphates and tryptophan or derivatives of tryptophan. The electronic nature of tryptophans, facilitating electrophilic attack at seven positions on the indole nucleus, allows Nature a plethora of opportunities to create structural diversity.[25] Prenylated tryptophans can be further divided into categories based on prenylation site and/or prenyl acceptor. The brevianamides (C2 of Trp-Pro dipeptides), the amauromines (C3 of Trp-Trp dipeptides), the asterriquinones (various sites of bis(indolyl)benzoquinones), and the ergot alkaloids (C4 of Trp) are examples of prenylated tryptophan secondary metabolites.[22]

Ergot alkaloids, secondary metabolites derived from 4-dimethylallyltryptophan (DMAT), are sorted into two categories – amide derivatives of lysergic acid and clavine alkaloids. Although both categories contain a tetracyclic ergoline ring system, the clavine alkaloids contain neither a carboxylic acid or amide derivative.[22] Lysergic acid diethylamide (LSD), a psychedelic drug, is unarguably the most well-known ergot

alkaloid. The structurally similar ergotamine is a lysergic acid derivative containing an amide-bonded peptide moiety. Given their structural similarities to the neurotransmitters noradrenaline, dopamine, and serotonin, ergot alkaloids are of great pharmaceutical interest as evidenced by the use of ergotamine as a medication for the treatment of acute migraines.[7] Therefore, much synthetic[25] and biosynthetic[22] effort has been devoted to understanding the chemistry, biology, and genetics of ergot alkaloids.

Prenyltransferases

Prenyltransferases are a family of enzymes that catalyze the alkylation of prenyl acceptors by allylic isoprenoid diphosphates. A variety of electron-rich nucleophiles are able to accept prenyl donation including alkenes (farnesyl diphosphate synthase[14]), thiols (protein farnesyltransferase and protein geranylgeranyltransferase[26]), amines (tRNA-dimethylallyl diphosphate transferase[27,28]), hydroxyl groups (geranylgeranylglyceryl phosphate synthase[29]), and aromatic rings (dimethylallyltrypotphan synthase[30]) (Figure 1.2).

Farnesyl diphosphate (FPP) synthase, perhaps the most studied prenyltransferase, is a member of the isoprenyl diphosphate synthase family.[31] FPP synthase catalyzes the synthesis of *E,E*-farnesyl diphosphate, a vital intermediate in the biosynthetic pathway of compounds such as cholesterol[32] and farnesylated and geranylgeranylated proteins.[33] To synthesize FPP, the enzyme catalyzes two 1'-4 couplings between a homoallylic acceptor and an allylic donor.[14,34] In the dissociative electrophilic alkylation mechanism, the dissociation of the diphosphate moiety from the allylic substrate DMAPP creates a resonance-stabilized carbocation that is electrophilically attacked by the 3,4-double bond of the homoallylic substrate IPP. Formation of a tertiary carbocation and subsequent

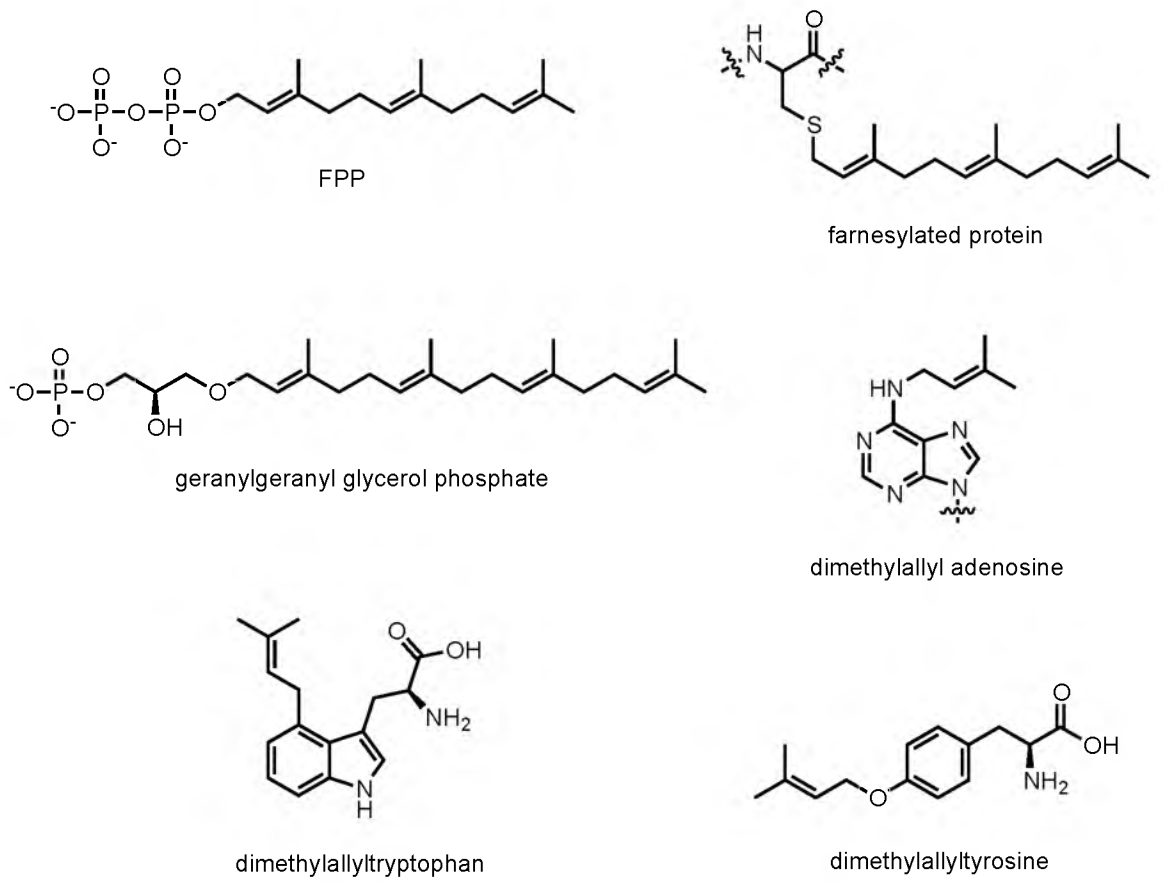


Figure 1.2. Chemical structures of representative prenyltransferase products highlighting the different types of prenyl acceptors.

elimination of a proton forms the C_{10} isoprenoid GPP. The newly formed GPP then acts as the allylic substrate for a second 1'-4 coupling with another molecule of IPP to form the C_{15} isoprenoid FPP. The crystal structures of avian[35] and *E. coli*[36] FPP synthases revealed a novel tertiary structure now known as the "isoprenoid synthase fold." The core is composed of 10 α-helices that surround a large active site cavity. This binding pocket includes two well-conserved DDXXD motifs, which coordinate the divalent cations needed for diphosphate stabilization and are important for catalysis.[37]

Protein prenyltransferases are another well-studied subgroup of prenyltransferases.[31] Many eukaryotic cellular proteins are post-translationally modified

with isoprenoid chains, thereby increasing their affinity for the membrane. These modified proteins include the Ras superfamily (Ras, Rab, Ral, Rac, and Rap), nuclear lamins, heterotrimeric G-proteins, and phosphodiesterases.[33,38,39] Protein farnesyltransferase (FTase) and protein geranylgeranyltransferase (GGTase) catalyze the divalent cation-dependent alkylation of sulfhydryls on C-terminal cysteines with FPP and GGPP, respectively. The conserved C-terminal CaaX motif is the site of prenylation, where "a" is typically an aliphatic amino acid and X is the amino acid that specifies whether farnesylation or geranylgeranylation occurs.[40]

Aromatic prenyltransferases are vital enzymes in the biosynthesis of secondary metabolites in plants, fungi, and bacteria.[41-44] Typically, they catalyze C-C bond formation between an aromatic carbon of the prenyl acceptor and C1' ("normal" prenylation) or C3' ("reverse" prenylation) of the prenyl donor. In some cases, bond formation can also occur between a heteroatom on an aromatic ring and C1' or C3' of the prenyl donor. Currently, aromatic prenyltransferases are grouped into four different categories – lipoquinone biosynthesis, plant secondary metabolism, "ABBA" prenyltransferases, and fungal indole prenyltransferases.

Aromatic prenyltransferases in lipoquinone biosynthesis catalyze the addition of long chain prenyl moieties to aromatic acceptors such as 1,4-dihydroxy-2-naphthoate (DHNA), homogentisate (HGA), or 4-hydroxybenzoate (4-HB). The DHNA octaprenyltransferase MenA from *E. coli* catalyzes simultaneous decarboxylation and C_{40} prenylation at C2 of DHNA.[45] HGA prenyltransferases also catalyze simultaneous decarboxylation and prenylation (C_{20}, C_{45}, or phytyl) of HGA, however, at different positions of the diphenol ring.[46,47] 4HB octaprenyltransferases, UbiA from *E. coli*[48] and

COQ2 from yeast,[49] catalyze the C_{40} prenylation at C3 of 4HB; these prenyltransferases do not catalyze decarboxylations. The UbiA family of prenyltransferases, found in humans, plants, bacteria, and yeast, are known for their isoprenoid substrate promiscuity, accepting short chain isoprenoids such as GPP and FPP.[48] The lipoquinone biosynthetic aromatic prenyltransferases include a conserved aspartate-rich motif, similar to FPP synthase.[41] Correspondingly, their activity is dependent on Mg^{2+} or other divalent cations. Being integral membrane proteins, no experimentally determined structures are available yet; a structural model of UbiA, however, was recently constructed.[50] Based on the soluble sesquiterpene cyclase 5-epi-aristolochene synthase from *Nicotiana tabacum*,[51] UbiA corresponded well to an all α-helical terpenoid synthase.[50]

One method to create structural diversity in plant secondary metabolism is the prenylation of aromatic compounds.[52] 4HB geranyltransferase and naringenin 8-dimethylallyltransferase (N8DT) are examples of flavonoid prenyltransferases.[53] The biosynthesis of shikonin, a plant naphthoquinone, requires the prenylation of 4HB. 4HB geranyltransferase, the first secondary metabolite prenyltransferase cloned from plants, catalyzes the geranylation at C3 of 4HB.[54] This reaction is identical to the UbiA family of enzymes, except 4HB geranyltransferase is specific for GPP. In the sophoraflavanone G biosynthetic pathway, N8DT catalyzes the condensation of naringenin and DMAPP. Both 4HB geranyltransferase and N8DT contain aspartate-rich motifs and are Mg^{2+} dependent.[41]

The ABBA[42,43] and fungal indole prenyltransferases[44] are distinctly different from the prenyltransferases described above. These enzymes are soluble, do not contain aspartate-rich regions, and are independent of divalent cations for activity (except for

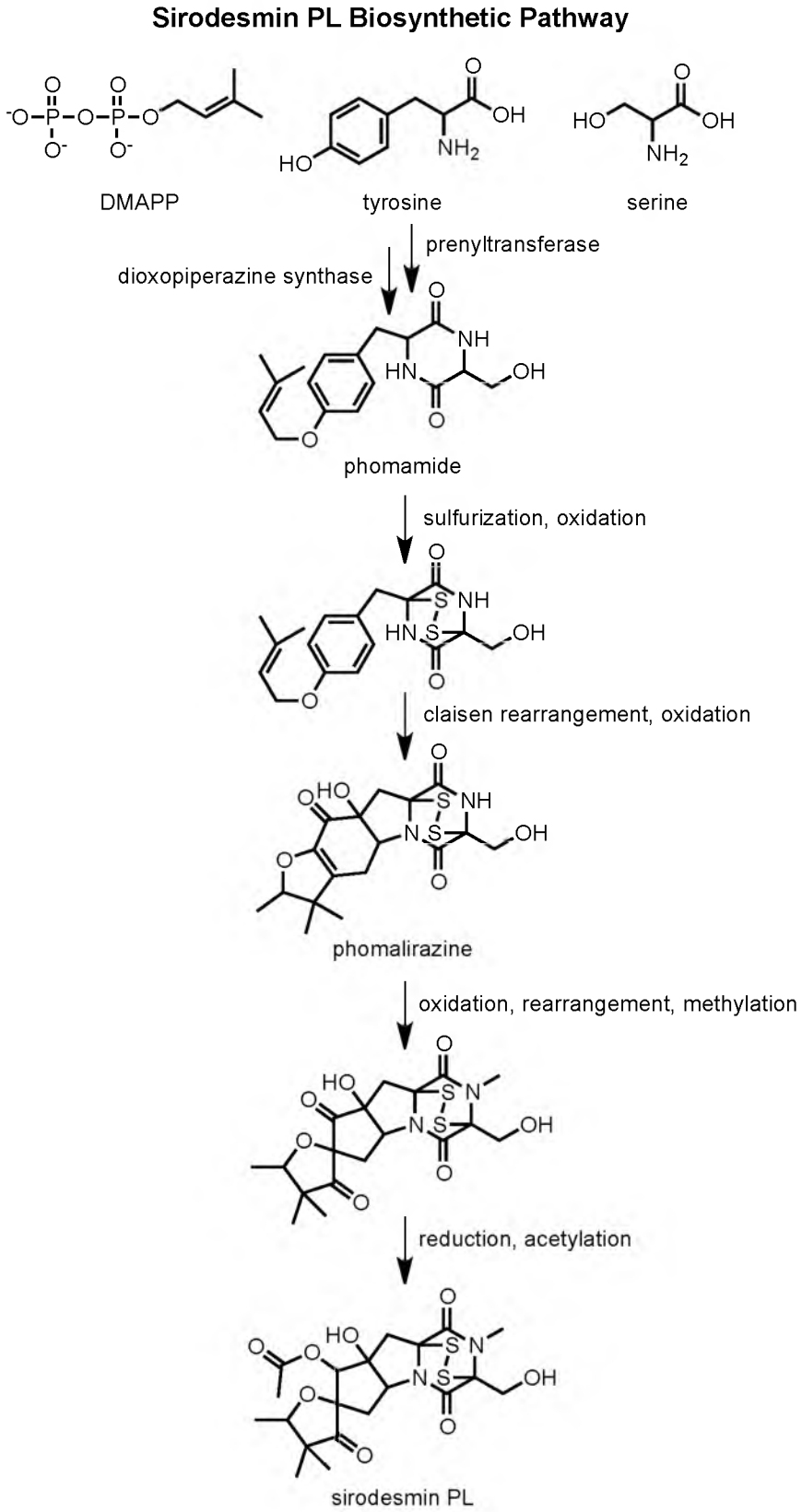
NphB). CloQ from *Streptomyces*, the first identified ABBA prenyltransferase, catalyzes the C3-prenylation of 4-hydroxyphenylpyruvate using DMAPP in the chlorobiocin biosynthetic pathway.[55] The subsequent crystallization of a sequence-related enzyme, NphB (formerly Orf2), revealed a previously unidentified antiparallel α-helix-β-sheet fold, later termed ABBA due to the α-β-β-α architecture.[42,56] Although the natural substrate is unknown, NphB catalyzes *C*-geranylations of phenolic substrates including 1,6-dihydroxynaphthalene.[56] CloQ was later crystallized and its structure was confirmed to be an ABBA fold.[57] Fnq26 from *Streptomyces cinnamonensis* DSM1042 was determined to catalyze both normal and reverse *C*-prenylations of flaviolin and normal *O*-prenylations of 4HB and 1,3-dihydroxynaphthalene.[58] Fnq26 was the first ABBA prenyltransferase seen to catalyze prenylations at C3' of the prenyl donor.

Fungal indole prenyltransferases catalyze alkylation of the indole ring in tryptophan, tryptophan-containing dipeptides, or indole-containing compounds with DMAPP to give the different naturally occurring carbon skeletons.[44] The dimethylallyl moiety can be attached at C1' or C3', further increasing structural diversity. Although fungal indole prenyltransferases contain little sequence similarity with bacterial ABBA prenyltransferases, biochemically, they are very similar. Fungal indole prenyltransferases are soluble, do not possess aspartate-rich motifs, are catalytically active in the absence of divalent cations, and are known to fold in the α-β-β-α pattern.[41,44,59-61] 4-Dimethylallyltryptophan synthase (4-DMATS) from *C. purpurea*, the first indole prenyltransferase purified, characterized, and cloned, catalyzes the normal prenylation of tryptophan at C4 by DMAPP.[62,63]

Dimethylallyltyrosine Synthase

The first known genes responsible for the biosynthesis of sirodesmin PL were reported in 2004.[64] Eighteen genes, spanning a 68-kilobase region, were found to encode the necessary enzymes for the production, regulation and self-resistance of sirodesmin PL in *Leptosphaeria maculans*. As known from early labeling experiments and intermediate analysis,[65-67] the biosynthesis of sirodesmin PL requires a nonribosomal peptide synthetase for the condensation of tyrosine and serine, a dimethylallyl transferase, and enzymes catalyzing a Claisen rearrangement, a sulfurization, an acetylation, a methylation, a reduction, and a number of oxidations (Scheme 1.3). Analysis of the eighteen genes revealed 16 potential ETP biosynthetic enzymes, one transcriptional regulator, and one ABC transporter assumed to control sirodesmin PL efflux and thus resistance.[64] The sixteen proposed biosynthetic enzymes include SirP, a two-module nonribomsomal peptide synthetase; SirT, a thioredoxin reductase for disulfide bond formation; SirH, an acetyl transferase; SirM and SirN, methyl transferases; SirB, SirC, and SirE, cytochrome P450 monooxygenases; SirO, an oxidoreductase; and SirD, a prenyltransferase.

The prenyltransferase gene, sirD, was found to be a homologue of a fungal DMATS from *Neotyphodium coenophialum*. Given the other members of the gene cluster and sirodesmin PL's structure, the SirD enzyme was hypothesized to prenylate the tyrosyl group of either free L-tyrosine or the cyclic dipeptide cyclo-L-Ser-L-Tyr (c-SY).[64] Pedras and Yu recently completed an isotopic feeding experiment, utilizing tritium-labeled biosynthetic intermediates to determine incorporation into sirodesmin PL.[68] L-tyrosine and *O*-dimethylallyl-L-tyrosine were efficiently incorporated into phomamide


Sirodesmin PL Biosynthetic Pathway
DMAPP
tyrosine
serine
prenyltransferase
dioxopiperazine synthase
phomamide
sulfurization, oxidation
claisen rearrangement, oxidation
phomalirazine
oxidation, rearrangement, methylation
reduction, acetylation
sirodesmin PL

Scheme 1.3

and sirodesmin PL, while c-SY was not. This suggested that the SirD-mediated prenylation of L-tyrosine is the first committed step in sirodesmin PL biosynthesis. Thus, SirD is a dimethylallyltyrosine synthase.

A BLAST[69] search reveals that SirD is a member of the aromatic prenyltransferase/DMATS superfamily. 7-DMATS from *A. fumigatus* is the closest relative studied to-date. SirD shares 35% identity and 52% similarity with 7-DMATS. Other relevant comparisons to sequence identity are: 31% with FgaPT2 from *A. fumigatus*, 27% with 4-DMATS from *C. purpurea*, and 27% with CTrpPT from *A. oryzae* DSM1147. A putative dimethylallyltyrosine synthase from *Sirodesmium diversum* is the nearest known relative of *L. maculans* SirD (87% identity), however the C-terminus is truncated by 165 residues compared to *L. maculans* SirD.

Dimethylallyltryptophan Synthase

Prenylated indole alkaloids constitute a large family of chemically diverse biologically active natural products.[22-24] Biosynthetically, the carbon skeletons of these compounds are constructed from prenyl diphosphates and tryptophan or derivatives of tryptophan. Collectively a family of enzymes, the dimethylallyltryptophan synthases (DMATS) catalyze alkylation of the indole ring in tryptophan or tryptophan-containing dipeptides at positions N1, C2, C3, C4, C5, C6, or C7 by DMAPP to give the different naturally occurring carbon skeletons.[63,70-76] The dimethylallyl moiety can be attached at C1' ("normal" prenylation) or C3' ("reverse" prenylation), further increasing structural diversity. For wild type DMATSs both normal and reverse prenylation is seen at N1[70,77] and C2.[71,78] Only reverse prenylation is seen at C3 in either the α- or β-configuration by the DMATSs AnaPT[72] or CdpC3PT,[73] respectively. Only normal prenylation is seen at

C4,[63,78] C5,[74] and C6.[75] 7-DMATS catalyzes normal prenylation at C7 of L-tryptophan[76] and MpnD catalyzes reverse prenylation at C7 of an indolactam.[79]

4-DMATS catalyzes the normal prenylation of tryptophan at C4 by DMAPP as the first committed step in ergot alkaloid biosynthesis.[30] The enzyme was the first indole prenyltransferase to be purified, characterized, and later available from expression clones and has been studied from *Claviceps purpurea*[62,63,80] and *Aspergillus fumigatus*.[59,78,81,82] BLAST alignments with 4-DMATS from *C. purpurea* and other DMATSs report the following sequence identities: 56% with FgaPT2 from *A. fumigatus*, 35% with FtmPT1 from *A. fumigatus*, 35% with CdpC3PT from *Neosartorya fischeri*, 31% with CdpNPT from *A. fumigatus*, 47% with 5-DMATS from *A. clavatus*, 23% with IptA from *Streptomyces* sp. SN-593, and 28% with 7-DMATS from *A. fumigatus*.[69]

Extensive work has been performed to understand the binding of substrates and the chemical mechanism of 4-DMATS. Initially, a random sequential binding mechanism was proposed for 4-DMATS in the presence of 20 mM Ca^{2+}.[30,83] A random sequential binding mechanism would allow either substrate, tryptophan or DMAPP, to bind first, followed by the binding of the second substrate to form an active complex. Subsequent studies including substrate binding, isotope trapping, and rapid quench experiments revealed the inability of DMAPP to bind in the absence of a divalent metal while tryptophan binds equally well in the presence or absence of metal.[84] It was also determined that very little DMAPP binds in the presence of metal. Since it is known that 4-DMATS is catalytically independent of divalent metal ions,[30,63,83,84] it is suggested a 4-DMATS:DMAPP complex is catalytically unproductive. Therefore, an ordered

sequential mechanism where tryptophan binds first, followed by DMAPP is proposed for prenylation to occur.[84]

The prenylation of indole rings has been of great interest to chemists and enzymologists alike. The delocalization of the lone electron pair on the nitrogen atom facilitates electrophilic substitution at positions 2-7 of the indole ring. However, the C4 position is normally understood to be an unfavorable position for substitution.[85] Consequently, the study of the C4 prenylation of tryptophan by 4-DMATS is of exceptional interest. Two mechanisms have been proposed for the chemical reaction performed by 4-DMATS: a direct electrophilic aromatic substitution at C4 or an electrophilic substitution at another position followed by a prenyl migration to C4.

Two labeled substrate studies support a dissociative electrophilic alkylation mechanism. Shibuya and coworkers verified an inversion of configuration at C1' of the dimethylallyl moiety while the allylic double bond geometry is preserved.[86] The authors proposed dissociation of the diphosphate moiety from C1' followed by subsequent C4 attack of C1' on the face opposite the pyrophosphate ion. Luk and Tanner observed scrambling of an ^{18}O-isotope label in [1-^{18}O]-DMAPP in a positional isotope exchange (PIX) experiment.[82] The scrambling of the ^{18}O label from the bridging position to a nonbridging position provides strong evidence for a dissociative mechanism where DMAPP dissociates into a stabilized allylic carbocation/pyrophosphate ion pair before electrophilic attack occurs.

Substrate analogue studies, specifically linear free energy relationships using 7-substituted tryptophan analogues and fluorinated DMAPP analogues, also support direct electrophilic aromatic substitution at the C4 position.[80] A Hammett plot of 7-substituted

tryptophan analogues generated a good linear correlation with ρ = -2.0. The authors concluded that although the value of -2.0 is at the low end for benzene derivatives, it is consistent with an electrophilic aromatic substitution mechanism.

Recently crystallized, 4-DMATS (FgaPT2) from *A. fumigatus* consists of a 10-stranded antiparallel β-barrel surrounded by α-helices (Figure 1.3).[59] To date, this "ABBA PT" fold is only seen for five other prenyltransferases - FtmPT1,[61] CdpNPT,[60] NphB,[56] CloQ,[57] and EpzP.[87] The crystal structure of FgaPT2 from *A. fumigatus* in complex with tryptophan and dimethylallyl *S*-thiolodiphosphate (DMASPP), a non-hydrolyzable DMAPP analogue,[88] revealed two active site amino acids thought to be important for catalysis.[59] A hydrogen bond between Glu89 and the indole N-H likely increases the electron density in the indole ring. Lys174 is situated near C4 and is a likely candidate for removing the C4 proton during rearomatization.

The mechanism currently proposed for prenylation by DMATSs is a dissociative electrophilic alkylation of the indole ring by DMAPP where cleavage of the carbon-oxygen bond in DMAPP gives a dimethylallyl cation-PP_i ion pair, with subsequent

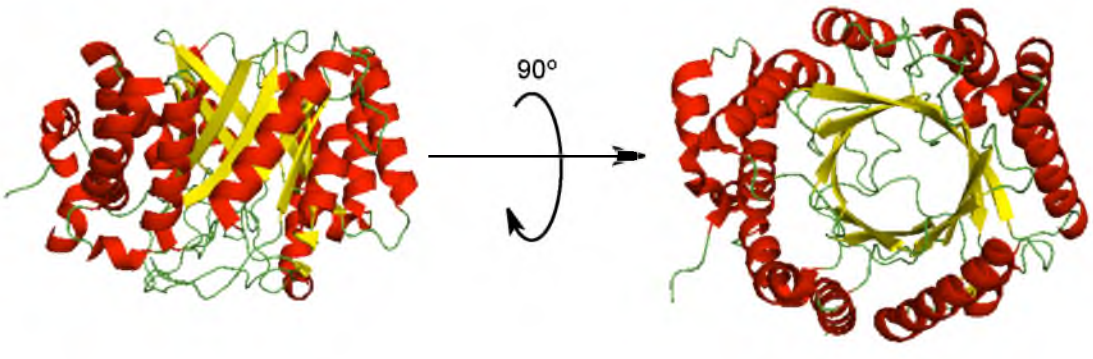


Figure 1.3. Cartoon representation of the 4-DMATS (FgaPT2) crystal structure depicting a side view (left) and top-down view (right). PDB: 3I4Z.

alkylation of the indole moiety, followed by loss of a proton to give the prenylated product. In the case of 4-DMATS, regiospecific alkylation of C4 of the indole ring by C1' of the allylic cation generates an arenium intermediate, which rearomatizes by deprotonation at C4 to produce dimethylallyltryptophan (DMAT) (Figure 1.4A). Recently, a K174A mutant of FgaPT2 gave a C3 reverse-prenylated hexahydropyrroloindole as the major product in addition to a small amount of DMAT.[89] The authors suggested a new mechanism where the initial alkylation is a reverse prenylation at C3 of the indole ring, followed by a Cope rearrangement to give the normal C4 prenylated arenium intermediate and deprotonation facilitated by Lys174 to produce DMAT (Figure 1.4B).

Dissertation Overview

The work presented in this dissertation describes the cloning and characterization of SirD, a dimethylallyltyrosine synthase, the synthesis of aromatic amino acid substrate analogues, and the aromatic substrate promiscuity of SirD and the dimethylallyltryptophan synthase 4-DMATS. Chapter 2 describes the chemical and enzymatic syntheses of substrates and substrate analogues for the tyrosine and tryptophan prenyltransferases. Chapter 3 describes the cloning and biochemical characterization of SirD, including the promiscuity studies involving tyrosine and tryptophan derivatives. Finally, Chapter 4 focuses on 4-DMATS and its ability to prenylate tryptophan derivatives, particularly analogues with substituents at the naturally alkylated C4 position.

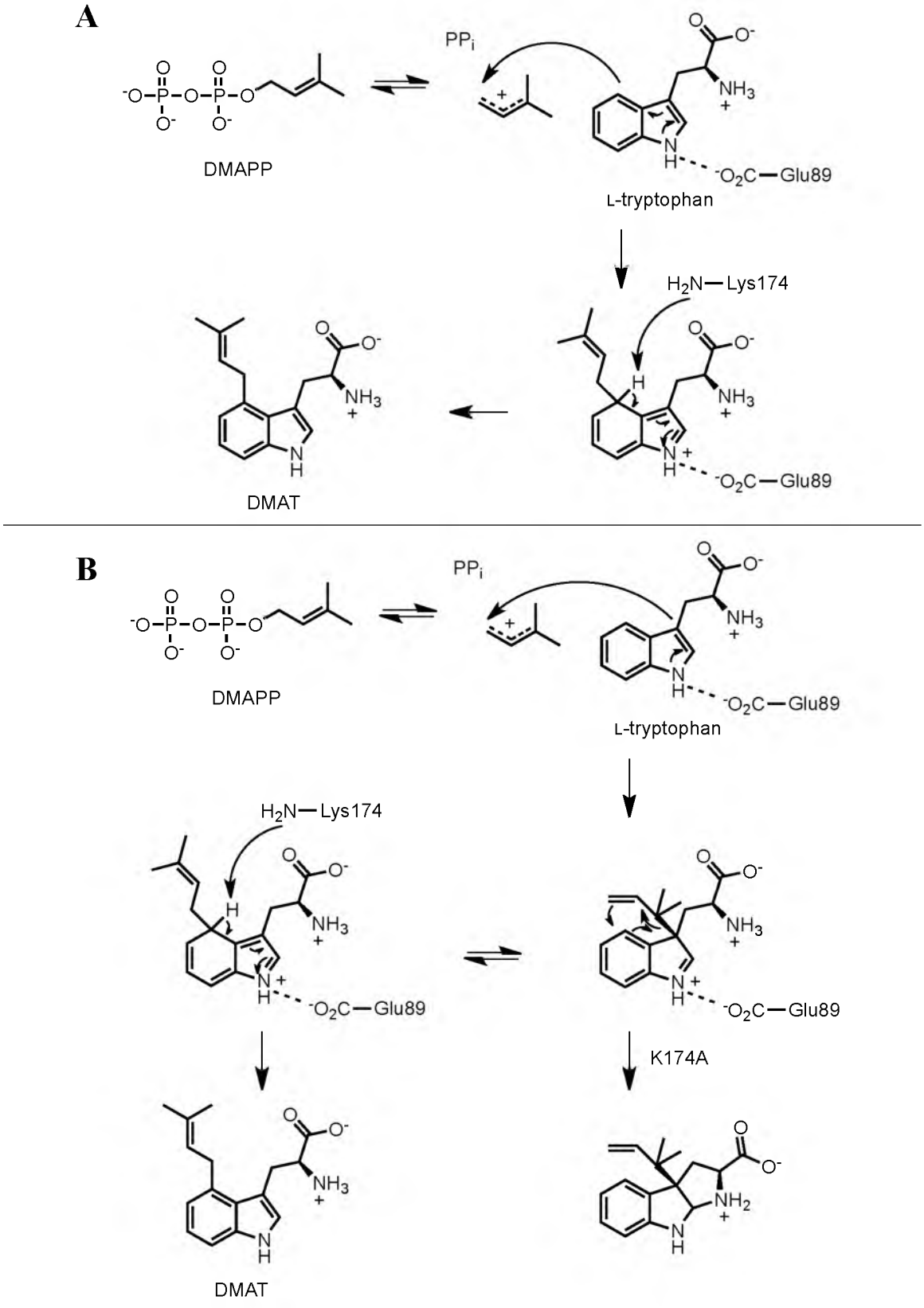


Figure 1.4. 4-DMATS proposed mechanisms. A, direct dissociative electrophilic alkylation; B, dissociative electrophilic alkylation and rearrangement.

CHAPTER 2

SYNTHESIS OF ENZYMATIC SUBSTRATES AND AROMATIC AMINO ACID ANALOGUES

Introduction

Vital to the study of promiscuous enzymes is the possession of a substrate library. Because SirD and DMATS prenylate aromatic amino acids, we were most interested in analogues of tyrosine, tryptophan, and phenylalanine. Certain analogues of interest, such as *O*-methyl-L-tyrosine and 4-amino-L-phenylalanine, are available commercially, while other analogues, such as 4-mercapto-L-phenylalanine and 4-vinyl-L-phenylalanine, must be synthesized.

A literature search for the synthesis of 4-mercapto-L-phenylalanine or L-thiotyrosine produces a two-step synthesis featuring the chlorosulfonation of L-phenylalanine followed by a Sn/HCl reduction.[90,91] The synthesis of a protected form of 4-vinyl-L-phenylalanine was also reported, however as an intermediate in the synthesis of 4-(phosphonomethyl)phenylalanine, an unnatural amino acid valuable for mimetic studies of protein-tyrosine kinases.[92] These two approaches were used as starting points for the synthesis of the 4-mercapto- and 4-vinyl-L-phenylalanine analogues.

The synthesis of indoles, which can be incorporated into tryptophans, has been extensively studied for well over 100 years. The most well known procedure is the Fischer indole synthesis, a condensation of phenylhydrazine with an aldehyde or ketone under acidic conditions to form indoles.[93] Other synthetic routes to indoles are also well documented: Reissert's condensation of nitrotoluene and diethyl oxalate followed by reductive cyclization;[94] Leimgruber's enamine formation from nitrotoluene and *N,N*-dimethylformamide dimethyl acetal and pyrrolidine followed by reductive cyclization;[95] and Sugasawa's chloroacetylation of anilines followed by reductive cyclization.[96] Once the desired indole is obtained, the amino acid moiety is built through the synthesis of

gramine.[97] Previously in the Poulter group, substituted tryptophans were synthesized using these established procedures.[98]

As noted above, the organic syntheses of indoles and tryptophans are well documented. However, not only are these synthetic pathways both step- and time-intensive, they also produce a racemic mixture of DL-tryptophans. Isolation of the pure L-tryptophans requires an additional step, such as an enzyme-catalyzed selective hydrolysis reaction[99] or purification by chiral chromatography. Alternatively, tryptophan synthase can be utilized, either as a cell extract or cell-free, to produce enantiomerically pure L-tryptophan analogues from a variety of indoles in one simple and convenient step.[100-102]

Tryptophan synthase is an $\alpha_2\beta_2$ tetramer that catalyzes the two final steps of tryptophan biosynthesis. The α-subunit catalyzes the formation of indole and glyceraldehyde-3-phosphate from indole-3-glycerol phosphate. The β-subunit then catalyzes the condensation of indole and L-serine to form L-tryptophan through a pyridoxal-5'-phosphate (PLP) dependent mechanism.[103] Conveniently, the β-subunit is active in the absence of the α-subunit allowing simple heterologous expression and purification of active tryptophan synthase β-subunit.

In this chapter, we report the chemical and enzymatic syntheses of substrates and substrate analogues for the tyrosine and tryptophan prenyltransferases.

Experimental Procedures

General

NMR spectra were recorded on a Varian Unity 300 MHz or Varian VXR 500 MHz and spectra were processed with MestReNova 7.1. All chemical shifts and

coupling constants are reported in parts per million (ppm) and hertz (Hz), respectively. ^{1}H chemical shifts are referenced relative to the following peaks: $CDCl_3$ at δ 7.24 ppm, DMSO-d_6 at 2.50 ppm (center peak), or DSS at 0.00 ppm. ^{13}C chemical shifts are referenced relative to the following peaks: $CDCl_3$ at δ 77.23 ppm (center peak), DMSO-d_6 at 39.51 (center peak), or DSS at 0.00 ppm. ^{19}F chemical shifts are referenced relative to trifluoroacetic acid at δ -76.55 ppm. ^{31}P chemical shifts are referenced relative to phosphoric acid in D_2O at δ 0 ppm. HPLC was conducted on a Water 2690 Separation Module equipped with a Microsorb MV™ C18 5 μm column and a Waters 996 photodiode array detector. Silica gel flash chromatrography was performed using ZEOprep 60 ECO silica gel (Zeochem). Thin layer chromatography was carried out on silica gel 60 glass-backed plates (Merck). Silica TLC plates were visualized under UV light or by staining with 10% w/v phosphomolybdic acid (Sigma) in ethanol or 1.5% w/v ninhydrin (Sigma) in n-butanol with 3% v/v acetic acid. Mass spectrometry was performed by Dr. James Muller at the Chemistry Department Mass Spectrometry Laboratory, University of Utah. Stock solutions of all diphosphates were prepared in 25 mM NH_4HCO_3. All diphosphate concentrations were determined using the phosphate analysis assay described by Lanzetta, et al.[104]

Materials

All compounds were obtained from Sigma-Aldrich unless otherwise noted. The following compounds were purchased from the companies indicated: triflic anhydride (Alfa Aesar); *N*-*t*-Boc-4-iodo-L-phenylalanine, 4-hydroxyindole, 7-hydroxyindole, 4-methoxyindole, 4-methylindole, 2-methylindole, and 4-azaindole (Chem-Impex International, Inc.); isopropyl β-D-1-thiogalactopyranoside (IPTG, Gold Biotechnology);

thionyl chloride and lithium chloride (Mallinckrodt/JT Baker); chloramphenicol (MP Biomedicals); Fmoc-OSu (Novabiochem); and cellulose (Whatman). Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) and deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. except sodium deuteroxide (NaOD) in D_2O, which was purchased from Aldrich.

Synthesis of Diphosphates

The isoprenoid diphosphates DMAPP and IPP were prepared according to literature procedures.[105,106]

Tris(tetra-*n*-butylammonium) Hydrogen Pyrophosphate (1)

J. Scott Lee generously supplied 10.06 g (11.16 mmol) of **1**.

3-Methyl-3-buten-1-yl *p*-Toluenesulfonate (Isopentenyl Tosylate, 2)

In a nitrogen atmosphere, 3-methyl-3-buten-1-ol (0.252 mL, 0.215 g, 2.50 mmol) was added to a solution of *p*-toluenesulfonyl chloride (0.55 g, 2.88 mmol) and 4-(*N*, *N*-dimethylamino)pyridine (0.36 g, 2.96 mmol) in DCM (12.5 mL). After allowing the reaction to stir at rt for 2.5 h, a 100 x volume excess of hexanes was added and the resulting precipitate was removed by gravity filtration. The filtrate was concentrated by rotary evaporation, diluted with diethyl ether, then filtered and concentrated to yield **2** (0.49 g, 2.04 mmol, 82%) as a colorless oil, which was used without further purification.

3-Methyl-3-buten-1-yl Diphosphate (Isopentenyl Diphosphate, 3)

In a nitrogen atmosphere, tosylate **2** (0.49 g, 2.04 mmol) in ACN (6.8 mL) was treated with **1** (5.53 g, 6.13 mmol) dissolved in ACN (6 mL). The reaction mixture was

allowed to stir at rt for 2 h. After removing the solvent by rotary evaporation, the residue was dissolved in a minimal amount of 1:49 IPA/25 mM NH_4HCO_3. The product was converted into its ammonium form by passing it through a DOWEX AG 50W-X8 (100 – 200 mesh) cation-exchange resin column (30 equiv.) and lyophilizing to dryness. The resulting white powder was dissolved in 0.1 M NH_4HCO_3, diluted with a mixture of ACN and IPA, vortexed until a white precipitate formed, and clarified by centrifugation (2000 rpm, 5 min). This process was repeated with the supernatant twice. The combined supernatants were concentrated *in vacuo* and lyophilized to yield a white powder. The white powder was then dissolved in a minimal amount of chromatography buffer and purified by cellulose flash chromatography using an elution system of 4.5:2.5:3 (v/v/v) IPA/ACN/0.1 M NH_4HCO_3 to yield **3** (0.11 g, 0.45 mmol, 22%) as a white solid; ^{1}H NMR (300 MHz, D_2O) δ 1.76 (s, 3H), 2.38 (t, J = 6.6 Hz, 2H), 4.04 (dd, J = 6.6, 13.2 Hz, 2H), ~4.8 (s, 2H, shielded by D_2O); ^{31}P NMR (121 MHz, D_2O) δ -7.19 (d, J = 17.8 Hz, 1P), -2.34 (d, J = 17.7 Hz, 1P). Except for the shielded C4 singlet, the NMR data matched the published data.[106]

1-Bromo-3-methyl-2-butene (Dimethylallyl Bromide, 4)

In a nitrogen atmosphere, phosphorus tribromide (0.43 mL, 1.23 g, 4.53 mmol) was added dropwise to an ice-cold solution of 3-methyl-2-buten-1-ol (1.05 mL, 0.90 g, 10.5 mmol) in pentane (15 mL). The reaction was stirred at 0 °C for 25 min before MeOH (1.4 mL, 3.3 x mol excess) was added. The solution was allowed to stir for another 5 min at 0 °C. After removal of the aqueous layer, the organic layer was filtered through a celite/$MgSO_4$ plug. The organic layer was then concentrated by rotary

evaporation in a 0 °C water bath to yield **4** (0.69 g, 4.63 mmol, 102%) as a brown oil, which was used without further purification.

3-Methyl-2-buten-1-yl Diphosphate (Dimethylallyl Diphosphate, 5)

In a nitrogen atmosphere, bromide **4** (0.69 g, 4.63 mmol) in ice-cold ACN (15.4 mL) was treated with **1** (8.7 g, 9.66 mmol) dissolved in ACN (19 mL). The reaction mixture was allowed to warm to rt and stirred for 2 h. After removing the solvent by rotary evaporation, the residue was converted into its ammonium form and purified according to the procedure described above for under the synthesis of **5** (IPP). **5** (0.52 g, 2.11 mmol, 46%) was obtained as a white solid; ^{1}H NMR (300 MHz, D_2O) δ 1.71 (s, 3H), 1.75 (s, 3H), 4.44 (t, J = 6.9 Hz, 2H), 5.44 (t, J = 6.9 Hz, 1H); ^{31}P NMR (121 MHz, D_2O) δ -11.74 (d, J = 17.8 Hz, 1P), -6.87 (d, J = 18.5 Hz, 1P). The NMR data matched the published data.[106]

Synthesis of *O*-Dimethylallyl-L-tyrosine

***N*-(*tert*-Butoxycarbonyl)-*O*-dimethylallyl-L-tyrosine (6)**

NaH (60% in mineral oil, 152 mg, 6.33 mmol) was washed with hexanes (3 x 5 mL) and added to a solution of *t*-Boc-L-tyrosine (613 mg, 2.18 mmol) in anhydrous THF (17.5 mL). After stirring at rt for 15 min, dimethylallyl bromide (**4**, 390 mg, 2.62 mmol) was added to the cloudy white mixture and stirred for 20 h. The resulting mixture was concentrated to a white foamy solid, diluted with 0.5 N NaOH (25 mL), and extracted with diethyl ether (3 x 20 mL). The aqueous layer was acidified to pH <2 and re-extracted with EtOAc (3 x 20 mL). The combined organic layer was dried with Na_2SO_4,

filtered, and concentrated to yield **6** (820 mg, 2.35 mmol, 108%). Compound **6** was used directly in the next step without any further purification.

***O*-Dimethylallyl-L-tyrosine (7)**

Compound **6** (820 mg, 2.35 mmol) was heated to 150 °C under nitrogen for 60 min. The orange solid was washed with ice-cold MeOH to give **7** (151 mg, 0.612 mmol, 26%) as a white solid. Compound **7** was further purified by silica gel chromatography with gradient elution using 0 – 20% H_2O in ACN; R_f 0.45 (ACN:H_2O 4:1); ^{1}H NMR (500 MHz, D_2O/NaOD, DSS) δ 1.68 (s, 3H), 1.73 (s, 3H), 2.68 (dd, J = 13.5, 23.0 Hz, 1H), 2.97 (dd, J = 8.5, 23.0 Hz, 1H), 3.42 (dd, J = 8.5, 13.0 Hz, 1H), 4.45 (d, J = 11.0 Hz, 2H), 5.43 (t, J = 10.5 Hz, 1H), 6.85 (d, J = 14.0 Hz, 2H), 7.14 (d, J = 14.0 Hz, 2H); ^{13}C NMR (125 MHz, D_2O/NaOD, DSS) δ 17.5, 25.1, 40.3, 57.5, 64.9, 114.9, 119.0, 130.5, 131.0, 139.6, 156.6, 182.2; UV (ACN:H_2O 1:3) λ_{max} 223.8, 273.6 nm; MS-ESI *m/z* 272.1 $[M + Na]^+$. The ^{1}H and ^{13}C NMR data matched the published data.[68]

Synthesis of 4-Mercapto-L-phenylalanine

***N*-(*tert*-Butoxycarbonyl)-4-(*S*-*tert*-butylthio)-L-phenylalanine (8)**

The tris(dibenzylideneacetone)dipalladium(0)-chloroform adduct (Pd_2dba_3·$CHCl_3$, 0.045 g, 0.044 mmol, 1.5 mol%) was added to DMF (15 mL) under nitrogen, followed by 1,1'-Bis(diphenylphosphino)ferrocene (DPPF, 0.111 g, 0.200 mmol, 6.6 mol %). The mixture was allowed to stir at rt for 10 min. A mixture of triethylamine (0.92 mL, 6.60 mmol) and *N*-*t*-Boc-4-iodo-L-phenylalanine (1.17 g, 3.00 mmol) in DMF (15 mL) was added dropwise and allowed to stir at rt for 10 min before the addition of 2-methyl-2-propanethiol (0.37 mL, 3.30 mmol). After stirring at 75 °C for 3 h, the reaction mixture

was concentrated *in vacuo*, diluted with ethyl acetate (40 mL) and water (10 mL), cooled in an ice bath and acidified to pH 2-3 with 50 mM citric acid. The aqueous phase was extracted with EtOAc (25 mL) and the combined organic layers were washed with brine (2 x 20 mL) and dried over Na_2SO_4. The concentrated crude product was purified by silica gel flash chromatography with gradient elution using 0 – 100% EtOAc in hexanes. The resulting yellow oil was re-dissolved in hexanes/EtOAc and filtered to remove a small amount of purple solid. The filtrate was concentrated to yield a thick, yellow oil of **8** (0.75 g, 2.12 mmol, 71%); R_f 0.41 (DCM:MeOH 9:1); ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.19 (s, 9H), 1.33 (s, 9H), 2.98 (dd, J = 6.9, 13.8 Hz, 1H), 3.16 (dd, J = 5.1, 13.8 Hz, 1H), 4.55 (q, J = 7.2, 13.2 Hz, 1H), 5.09 (d, J = 8.1 Hz, 1H), 7.09 (d, J = 7.8 Hz, 2H), 7.38 (d, J = 7.8 Hz, 2H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 28.1, 28.3, 31.0, 37.7, 45.9, 54.2, 80.2, 129.6, 131.2, 137.1, 137.6, 155.4, 175.1. The ^{1}H NMR matched the literature in DMSO-d_6, however no ^{13}C NMR was previously reported.[107]

Bis-(4-mercapto-L-phenylalanine) Disulfide (9)

Concentrated HCl (8 mL) was added to **8** (220 mg, 0.62 mmol) and was heated at reflux overnight. After cooling to rt, the solution was diluted with H_2O (20 mL) and washed with EtOAc (3 x 20 mL). The resulting aqueous layer was lyophilized to yield a brown solid of **9** (87 mg, 0.22 mmol, 71%); R_f 0.24 (ACN:H_2O 4:1), 0.34 (thiol) and 0.48 (disulfide) (n-BuOH:AcOH:H_2O:pyridine 15:3:10:6); ^{1}H NMR (500 MHz, DMSO-d_6) δ 3.12 (d, 6.5 Hz, 4H), 4.15 (t, 6.5 Hz, 2H), 7.31 (d, 8.5 Hz, 4H), 7.49 (d, 8.0 Hz, 4H), 8.44 (br, 4H); ^{13}C NMR (125 MHz, DMSO-d_6) δ 35.1, 53.0, 127.4, 130.7, 134.6, 134.7, 170.2; UV (ACN:H_2O 1:3) λ_{max} 206.2, 242.8 nm; MS-ESI *m/z* 198.0 $[Monomer + H]^+$,

393.2 $[M + H]^+$, 415.3 $[M + Na]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{18}H_{21}N_2O_4S_2$ 393.0943, found 393.0945.

Synthesis of 4-Vinyl-L-phenylalanine

L-Tyrosine Methyl Ester (10)

Thionyl chloride (4 mL, 55.0 mmol) was added dropwise to a suspension of L-tyrosine (5.45 g, 30.1 mmol) in MeOH (100 mL, 247 mmol) at 0 °C. The mixture was allowed to warm to rt and stirred overnight. After removing the solvent by rotary evaporation, the residue was washed with diethyl ether (2x50 mL). Compound **10** was obtained as a white solid (5.7 g, 29.2 mmol, 97%); ^{1}H NMR (300 MHz, D_2O) δ 3.13 (dd, *J* = 7.2, 14.7 Hz, 1H), 3.24 (dd, *J* = 5.7, 14.7 Hz, 1H), 3.82 (s, 3H), 4.36 (dd, *J* = 5.7, 7.2 Hz, 1H), 6.87 (d, *J* = 8.7 Hz, 2H), 7.13 (d, *J* = 8.7 Hz, 2H). The ^{1}H NMR data matched the published data,[108] except in my spectrum the α proton (4.36) is resolved from the methyl group (3.82).

N-(9-Fluorenylmethoxycarbonyl)-L-tyrosine Methyl Ester (11)

A solution of *N*-(9-fluorenylmethoxycarbonyl)succinimide (4.89 g, 14.5 mmol) in 1,4-dioxane (30 mL) was added to a mixture of **10** (2.50 g, 12.8 mmol) and $NaHCO_3$ (3.11 g, 37.0 mmol) in H_2O (30 mL). After stirring for 23 h at rt, the mixture was acidified to pH 2-3 with 1 N HCl at 0 °C, extracted with EtOAc (3 x 25 mL), dried over Na_2SO_4, and concentrated *in vacuo*. The crude product was purified by silica gel flash chromatography with gradient elution using 0 – 40% EtOAc in hexanes to yield **11** (4.8 g, 90%) as a colorless sticky solid; ^{1}H NMR (300 MHz, $CDCl_3$) δ 2.93-3.09 (m, 2H), 3.70 (s, 3H), 4.18 (t, *J* = 6.9 Hz, 1H), 4.27-4.46 (m, 2H), 4.57-4.67 (m, 1H), 5.35 (d, 8.4

Hz, 1H), 6.72 (d, J = 8.4 Hz, 2H), 6.91 (d, J = 8.4 Hz, 2H), 7.25-7.43 (m, 4H), 7.51-7.57 (m, 2H), 7.71-7.78 (m, 2H). The ^{1}H NMR matched the published data.[109]

_N_-9-Fluorenylmethoxycarbonyl-4-((trifluoromethane-sulfonyl)oxy)-L-phenylalanine Methyl Ester (12)

Trifluoromethanesulfonic anhydride (1.9 mL, 11.3 mmol) was added dropwise to an ice-cold solution of **11** (4.1 g, 9.82 mmol) and DIPEA (8.6 mL, 49.4 mmol) in DCM (20 mL). Stirring was continued at 0 °C for 2.5 h. The resulting mixture was diluted with H_2O (30 mL) and DCM (50 mL) and washed sequentially with 10% $NaHCO_3$ (25 mL), H_2O (30 mL), 10% citric acid (2 x 30 mL), and H_2O (30 mL). The organic layer was dried over Na_2SO_4 and concentrated *in vacuo*. The crude product was purified by silica gel flash chromatography using an elution gradient of 0 – 30% EtOAc in hexanes to give **12** (3.0 g, 5.46 mmol, 56%) as a white solid; R_f 0.43 (hexanes:EtOAc 4:1); ^{1}H NMR (300 MHz, $CDCl_3$) δ 3.07 (dd, J = 6.0, 13.8 Hz, 1H), 3.16 (dd, J = 6.0, 14.1 Hz, 1H), 3.70 (s, 3H), 4.19 (t, J = 6.6 Hz, 1H), 4.38 (dd, 6.3, 10.5 Hz, 1H), 4.48 (dd, 6.9, 10.8 Hz, 1H), 4.64 (dd, J = 6.0, 13.8 Hz, 1H), 5.27 (d, J = 7.8 Hz, 1H), 7.12 (d, J = 8.7 Hz, 2H), 7.17 (d, J = 8.7 Hz, 2H), 7.30 (t, J = 7.5 Hz, 2H), 7.40 (t, J = 7.5 Hz, 2H), 7.55 (d, J = 7.5 Hz, 2H), 7.76 (d, J = 7.5 Hz, 2H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 37.8, 47.4, 52.7, 54.8, 67.0, 120.2, 120.3, 121.6, 125.1, 125.2, 127.3, 128.0, 131.3, 136.7, 141.6, 143.8, 143.9, 148.8, 155.6, 171.6; ^{19}F NMR (282 MHz, $CDCl_3$, TFA) δ 73.8; MS-ESI *m/z* 572.1 $[M + Na]^+$, HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{26}H_{22}NO_7NaSF_3$ 572.0967, found 572.0970. The ^{1}H NMR matched the published data; however, the ^{13}C NMR had a few differences in the aromatic region and no ^{19}F NMR data were previously reported.[110]

4-Vinyl-L-phenylalanine Methyl Ester (13)

Under a nitrogen atmosphere, tributyl(vinyl)tin (267 μL, 0.92 mmol) was added to a solution of flame-dried LiCl (270 mg, 6.35 mmol), $PdCl_2(PPh_3)_2$ (12.5 mg, 0.018 mmol), and triflate **12** (0.5 g, 0.91 mmol) in DMF (8 mL). The yellow-orange solution was then warmed to 90 °C. After stirring for 23 h, the reaction mixture was poured into ice-cold H_2O (15 mL) and extracted with EtOAc (2 x 10 mL). The combined organic extracts were sequentially washed with H_2O (2 x 10 mL) and brine (2 x 10 mL), dried over Na_2SO_4, and concentrated *in vacuo* to give a yellow oil. This oil was partially purified by silica gel flash chromatography with gradient elution using 0 – 100% EtOAc in hexanes, followed by 10% MeOH in EtOAc. Impure 4-vinyl-L-phenylalanine methyl ester (**13**, 0.13 g, 0.63 mmol, 70%) was recovered from the MeOH fractions; R_f 0.55 (EtOAc:MeOH 4:1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.83 (dd, J = 8.0, 13.5 Hz, 1H), 3.05 (dd, J = 5.5, 13.5 Hz, 1H), 3.61-3.77 (m, 4H), 5.20 (d, J = 11.0 Hz, 1H), 5.70 (d, J = 17.5 Hz, 1H), 6.67 (dd, J = 18.0, 11.0 Hz, 1H), 7.13 (d, J = 8.0 Hz, 2H), 7.33 (d, J = 8.0 Hz, 2H); MS-ESI *m/z* 206.1 $[M + H]^+$, 228.1 $[M + Na]^+$.

***N*-9-Fluorenylmethoxycarbonyl-4-vinyl-L-phenylalanine Methyl Ester (14)**

In an attempt to better purify the vinyl amino acid, impure **13** (0.13 g, 0.63 mmol) was reprotected with Fmoc in a procedure similar to the synthesis of **11**, using $NaHCO_3$ (0.16 g, 1.91 mmol) and *N*-(9-fluorenylmethoxycarbonyl)succinimide (0.26 g, 0.77 mmol). Silica gel flash chromatography using an elution gradient of 0 – 20% EtOAc in hexanes afforded **14** (0.15 g, 0.35 mmol, 55%). R_f 0.22 (hexanes:EtOAc 4:1), 0.60 (hexanes:EtOAc 3:2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.07 (dd, J = 6.0, 13.5 Hz, 1H),

3.12 (dd, J = 6.0, 14.0 Hz, 1H) 3.72 (s, 3H), 4.20 (t, J = 7.0 Hz, 1H), 4.34 (dd, J = 6.5, 10.5 Hz, 1H), 4.42 (dd, J = 7.0, 10.5 Hz, 1H), 4.65 (dd, J = 5.5, 13.5 Hz, 1H), 5.19-5.26 (m, 2H), 5.71 (d, J = 18.0 Hz, 1H), 6.67 (dd, J = 11.0, 17.5 Hz, 1H), 7.03 (d, J = 7.5 Hz, 2H), 7.27-7.33 (m, 4H), 7.39 (t, J = 7.5 Hz, 2H), 7.55 (t, J = 7.0 Hz, 2H), 7.75 (d, J = 7.5 Hz, 2H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 38.1, 47.4, 52.6, 54.9, 67.2, 114.0, 120.1, 120.2, 120.3, 124.9, 125.2, 125.3, 126.7, 127.2, 127.3, 127.8, 127.9, 129.7, 135.5, 136.6, 136.7, 141.5, 141.8, 143.9, 144.1, 155.7, 172.1; MS-ESI *m/z* 450.1 $[M + Na]^+$, HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{27}H_{25}NO_4Na$ 450.1681, found 450.1688.

4-Vinyl-L-phenylalanine (15)

Protected vinyl amino acid **14** (0.13 g, 0.30 mmol) was dissolved in a 1:1 mixture of 1 N NaOH (5x mol excess)/THF and allowed to stir at rt for 42 h. The reaction mixture was washed with diethyl ether (3 x 20 mL) and the aqueous layer was subsequently lyophilized to yield an off-white solid (0.056 g, 0.29 mmol, 96%). Product **15** was dissolved in H_2O and purified by C18 RP-HPLC using a linear elution gradient of 100% H_2O to 100% ACN in a total volume of 30 mL at a flow rate of 1.0 mL/min. Fractions with retention times 9-11 min were collected and lyophilized to afford a white solid; ^{1}H NMR (300 MHz, D_2O/NaOD, DSS) δ 2.92 (dd, J = 7.8, 14.1 Hz, 1H), 3.06 (dd, J = 5.7, 14.1 Hz, 1H), 3.64 (dd, J = 5.7, 7.5 Hz, 1H), 5.28 (d, J = 11.1 Hz, 1H), 5.83 (d, J = 17.7 Hz, 1H), 6.78 (dd, J = 10.8, 17.7 Hz, 1H), 7.26 (d, J = 8.4 Hz, 2H), 7.47 (d, J = 8.4 Hz, 2H); ^{13}C NMR (75 MHz, D_2O/NaOD, DSS) δ 39.2, 57.1, 114.1, 126.5, 129.9, 136.2, 136.5, 137.2, 179.9; UV (ACN:H_2O 1:3) λ_{max} 202.6, 249.9 nm; MS-ESI *m/z* 214.1 $[M + Na]^+$, 236.1 $[M + 2Na^+ - H^+]^+$; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{11}H_{13}NO_2Na$ 214.0844, found 214.0838.

Expression of the *T. maritime* Tryptophan Synthase β-Subunit

Bacterial Strains, Plasmids and Culture Conditions

tmTrpB, an *E. coli* BL21-(DE3)-RIPL strain containing the pET28a-tmTrpB1 vector was obtained from Reinhard Sterner. tmTrpB1 encodes the β-subunit of tryptophan synthase from the hyperthermophilic bacterium *Thermotoga maritime*.[111] Expression strains were grown in liquid LB broth or on solid LB-agar medium at 37 °C. Kanamycin (35 μg mL^{-1}) and chloramphenicol (34 μg mL^{-1}) were used to select vector-containing *E. coli* strains.

Overproduction and Purification of Tryptophan Synthase

As advised by the Sterner lab, *E. coli* BL21-DE3-RIPL containing pET28a-tmTrpB1 was cultivated in 1 L LB broth containing kanamycin and chloramphenicol at 37 °C with shaking at 225 rpm to an OD_{600} of 0.5 – 1.0. The culture was then cooled to 20 °C and allowed to grow for 1 h. Overexpression of the β-subunit of tryptophan synthase was induced with IPTG to a final concentration of 0.5 mM. After inducing for 18 h, cells were harvested by centrifugation (2,700 x *g*, 15 min, 4 °C). Pelleted cells were resuspended in 25 mL lysis buffer (100 mM K_2HPO_4, 300 mM KCl, 10 mM imidazole, 40 μM PLP, pH 7.5) and sonicated (Branson Sonifier 350) for 4 x 30 sec at 4 °C. The lysate was clarified by centrifugation (23,400 x *g*, 15 min, 4 °C). The resulting soluble lysate was heated to 75 °C for 20 min and again clarified by centrifugation (23,400 x *g*, 15 min, 4 °C). Purification of the recombinant protein was achieved through nickel affinity chromatography (GE Healthcare HisTrap HP) following the manufacturer's instructions. Tryptophan synthase was eluted with elution buffer (100 mM K_2HPO_4, 300 mM KCl, 1 M imidazole, pH 7.5) and pure protein was dialyzed 3 x

against 4 L of 100 mM K_2HPO_4 buffer, pH 7.5. The protein was then divided into 1 mL portions, flash frozen with liquid nitrogen and stored at -80 °C until use.

Synthesis of Indole Analogues

4-((Trifluoromethanesulfonyl)oxy)-indole (16)

Under a nitrogen atmosphere, trifluoromethanesulfonic anhydride (0.76 mL, 4.52 mmol) was added dropwise to an ice-cold solution of 4-hydroxyindole (500 mg, 3.76 mmol) and 2,6-lutidene (0.57 mL, 4.89 mmol) in DCM (5.5 mL). While stirring, the reaction was held at 0 °C for 1 h. The resulting yellow-brown mixture was poured into H_2O (20 mL). The aqueous layer was extracted with EtOAc (2 x 20 mL) and the combined organic fractions were washed with H_2O (20 mL) and brine (20 mL). The organic layer was then dried over Na_2SO_4 and concentrated *in vacuo*. The crude product **16** was isolated as an oil and used without further purification; R_f 0.57 (hexanes:EtOAc 3:2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.61-6.65 (m, 1H), 7.04 (d, J = 13.0 Hz, 1H), 7.15 (t, J = 13.0 Hz, 1H), 7.21 (dd, J = 4.5 Hz, 5.5 Hz, 1H), 7.34 (dt, J = 1.5 Hz, 14.0 Hz, 1H), 8.58 (br, 1H); ^{19}F NMR (470 MHz, $CDCl_3$) δ 73.7. The ^{1}H NMR matched the published data; however, no ^{19}F NMR data were previously reported.[112]

4-(*S*-*tert*-Butylthio)-indole (17)

Tris(dibenzylideneacetone)dipalladium(0)-chloroform adduct (Pd_2dba_3·$CHCl_3$, 0.056 g, 0.055 mmol, 1.5 mol%) was first added to DMF (18.7 mL) under nitrogen. 1,1'-Bis(diphenylphosphino)ferrocene (DPPF, 0.138 g, 0.249 mmol, 6.6 mol%) was then added to the purple solution and allowed to stir at rt for 10 min. A mixture of triethylamine (1.14 mL, 8.235 mmol) and triflate **16** (crude, assumed 3.76 mmol) in DMF

(18.7 mL) was added dropwise and allowed to stir at rt for 10 min before the addition of 2-methyl-2-propanethiol (0.464 mL, 4.11 mmol). After stirring at 75 °C for 18 h, the reaction mixture was diluted with EtOAc (50 mL) and H_2O (20 mL), cooled in an ice bath and acidified to pH 2-3 with 50 mM citric acid. The aqueous phase was extracted with EtOAc (2 x 25 mL) and the combined organic layers were washed with H_2O (2 x 25 mL) and brine (2 x 25 mL) and dried over Na_2SO_4. The organic layer was concentrated by rotary evaporation to a red-brown oil and purified by silica gel flash chromatography with gradient elution using 0 – 20% EtOAc in hexanes to give **17** (0.60 g, 2.92 mmol, 78%) as a yellow solid; R_f 0.45 (EtOAc:MeOH 4:1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 1.31 (s, 9H), 6.80-6.83 (m, 1H), 7.14 (t, J = 7.5 Hz, 1H), 7.22 (t, J = 3.0 Hz, 1H), 7.31 (d, J = 7.0 Hz, 1H), 7.39 (d, J = 8.0 Hz, 1H), 8.20 (br, 1H); ^{13}C (125 MHz, $CDCl_3$) δ 31.7, 47.4, 104.1, 112.1, 121.9, 124.5, 124.6, 130.6, 133.8, 135.9; GCMS *m/z* 205 [M]; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{12}H_{16}NS$ 206.1003, found 206.1007.

4-Mercaptoindole (18)

Concentrated HCl (6 mL) was added to **17** (180 mg, 0.88 mmol) and the resulting tan solution with brown precipitate was heated at reflux overnight. After cooling to rt, the solution was diluted with MeOH (2 mL) and H_2O (20 mL) and extracted with EtOAc (3 x 20 mL). The resulting organic layer was concentrated *in vacuo* and purified by silica gel flash chromatography with gradient elution using 0 – 40% EtOAc in hexanes to give **18** as a brown oil (40 mg, 0.27 mmol, 31%); R_f 0.53 (Hex:EtOAc 3:2); ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.02 (br, 1H), 6.62 (d, J = 7.5 Hz, 1H), 7.15 (t, J = 7.8 Hz, 1H), 7.26 (d, J = 5.7 Hz, 1H), 7.29-7.36 (m, 2H); ^{13}C (75 MHz, $CDCl_3$) δ 109.1, 113.2, 119.4, 124.6, 125.6, 128.4, 141.3, 141.7.

4-Vinylindole (19)

Under a nitrogen atmosphere, tributyl(vinyl)tin (0.922 mL, 3.18 mmol) was added to a solution of flame-dried LiCl (1.12 g, 26.36 mmol), $PdCl_2(PPh_3)_2$ (52.7 mg, 0.075 mmol), and triflate **16** (crude, assumed 3.76 mmol) in DMF (21 mL). The solution was then warmed to 80 °C. After stirring for 20 h, the reaction mixture was poured into ice-cold H_2O (80 mL) and extracted with EtOAc (3 x 25 mL). The combined organic extracts were washed with brine (2 x 25 mL), dried over Na_2SO_4, and concentrated *in vacuo* to a brown oil. The crude product was purified by silica gel flash chromatography with gradient elution using 0 – 40% EtOAc in hexanes to give **19** as a light brown liquid (0.29 g, 2.03 mmol, 64%); R_f 0.85 (EtOAc:MeOH 3:2); ^{1}H NMR (500 MHz, $CDCl_3$) δ 5.43 (dd, J = 2.0, 18.5 Hz, 1H), 5.97 (dd, J = 2.0, 29.5 Hz, 1H), 6.76-6.81 (m, 1H), 7.11-7.34 (m, 5H), 8.19 (br, 1H); ^{13}C (125 MHz, $CDCl_3$) δ 101.0, 110.8, 114.7, 117.4, 122.1, 124.7, 126.1, 130.1, 135.5, 136.3; MS-ESI *m/z* 144.1 $[M + H]^+$. The ^{1}H NMR matched the published data, however no ^{13}C NMR data was previously reported.[113]

Synthesis of L-Tryptophan Analogues

General Tryptophan Analogue Synthesis Method

Typically, indole analogue (160 mg) and L-serine (250 mg, 2.38 mmol) were added to 25 mL of Buffer TS (100 mM K_2HPO_4, pH 7.5, 180 mM KCl, 120 μM PLP) containing 1 – 8 mg of tryptophan synthase. For most analogues (4-OH, 7-OH, 4-OMe, 4-NH_2, 4-Me, 7-Me, 4-Vinyl, DIT), the reactions were carried out under nitrogen. The reactions were heated to 80 °C and incubated 16 - 90 h. After incubation, the reaction mixture was cooled to rt and the enzyme was removed by filtration using an Amicon Ultra 10 kDa MWCO filter. After concentrating *in vacuo*, the resulting residue was

purified by silica gel flash chromatography with gradient elution using 0 – 20 % H_2O in ACN. All purified compounds were stored under N_2 at -20 °C unless otherwise noted.

2-Methyl-L-tryptophan (20)

2-Methylindole (160 mg, 1.22 mmol) was incubated with 3.1 mg of tryptophan synthase as described above for 24 h. Column chromatography yielded **20** (140 mg, 0.64 mmol, 53%) as a pink solid; R_f 0.37 (ACN:H_2O 4:1); ^{1}H NMR (500 MHz, D_2O, DSS) δ 2.40 (s, 3H), 3.20 (dd, *J* = 8.5, 15.5 Hz 1H), 3.41 (dd, *J* = 5.0, 15.5 Hz, 1H), 4.00 (dd, *J* = 5.5, 8.5 Hz, 1H), 7.14 (dt, *J* = 1.0, 8.0 Hz, 1H), 7.20 (dt, *J* = 1.5, 7.5 Hz, 1H), 7.43 (d, *J* = 8.0 Hz, 1H), 7.61 (d, *J* = 7.5 Hz, 1H); UV (ACN:H_2O 1:3) λ_{max} 221.5, 227.2 nm; MS-ESI *m/z* 241.1 $[M + Na]^+$, 257.1 $[M + K]^+$; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{12}H_{14}N_2O_2Na$ 241.0953, found 241.0964. The ^{1}H NMR matched the literature.[100]

4-Amino-L-tryptophan (21)

4-Aminoindole (160 mg, 1.21 mmol) was incubated with 2 mg of tryptophan synthase as described above for 16 h. Column chromatography yielded **21** (220 mg, 1.00 mmol, 83%) as a gray solid; R_f 0.34 (ACN:H_2O 4:1); ^{1}H NMR (500 MHz, D_2O) δ 3.38 (dd, *J* = 8.0, 16.0 Hz, 1H), 3.62 (dd, *J* = 5.0, 16.0 Hz, 1H), 4.02 (dd, *J* = 5.0, 8.0 Hz, 1H), 6.57 (t, *J* = 4.5 Hz, 1H), 7.08-7.10 (m, 2H), 7.20 (s, 1H); UV (ACN:H_2O 1:3) λ_{max} 221.5, 272.4 nm; MS-ESI *m/z* 220.1 $[M + Na]^+$, 242.2 $[M + Na]^+$, 258.1 $[M + K]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{11}H_{14}N_3O_2$ 220.1086, found 220.1078. The ^{1}H NMR matched the literature, except the chemical shift of one aromatic proton.[101]

4-Hydroxy-L-tryptophan (22)

4-Hydroxyindole (160 mg, 1.20 mmol) was incubated with 8 mg of tryptophan synthase as described above for 90 h. Column chromatography yielded **22** (200 mg, 0.91 mmol, 76%) as light brown-gray crystals, which were stored under argon at -80 °C; R_f 0.42 (ACN:H_2O 4:1); ^{1}H NMR (300 MHz, D_2O) δ 3.07 (dd, J = 8.4, 15.0 Hz, 1H), 3.47 (ddd, J = 0.9, 4.5, 15.0 Hz, 1H), 3.95 (dd, 4.5, 8.1 Hz, 1H), 6.38 (dd, J = 3.6, 4.8 Hz, 1H), 6.89-6.91 (m, 2H), 6.97 (s, 1H); UV (ACN:H_2O 1:3) λ_{max} 219.1, 265.3, 280.7, 290.3 nm; MS-ESI *m/z* 243.1 $[M + Na]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{11}H_{13}N_2O_3$ 221.0926, found 221.0934. The ^{1}H NMR matched the literature.[114]

4-Mercapto-L-tryptophan (23)

4-Mercaptoindole (**18**, 20 mg, 0.14 mmol) was incubated with 1 mg of tryptophan synthase in Buffer TS containing 20 mM 2-mercaptoethanol (βME) as described above for 24 h. Column chromatography yielded **23** as a light yellow solid; R_f 0.41 (ACN:H_2O 4:1); UV (ACN:H_2O 1:3) λ_{max} 213.2, 271.2, 317.7 nm; MS-ESI *m/z* 237.1 $[M + H]^+$, 259.1 $[M + Na]^+$; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{11}H_{12}N_2O_2NaS$ 259.0517, found 259.0524.

4-Methoxy-L-tryptophan (24)

4-Methoxyindole (160 mg, 1.09 mmol) was incubated with 8 mg of tryptophan synthase as described above for 24 h. Column chromatography yielded **24** (120 mg, 0.51 mmol, 47%) as a light yellow solid; R_f 0.44 (ACN:H_2O 4:1); ^{1}H NMR (500 MHz, D_2O) δ 3.23 (dd, J = 8.5, 14.5 Hz 1H), 3.67 (dd, J = 4.5, 14.5 Hz, 1H), 4.01 (s, 3H), 4.14 (dd, J = 4.5, 8.5 Hz, 1H), 6.70 (d, J = 7.0 Hz, 1H), 7.17 (d, J = 8.0 Hz, 1H), 7.18 (s, 1H), 7.22 (t, J

= 7.5 Hz); ^{13}C NMR (125 MHz, D_2O) δ 28.0, 55.5, 56.5, 100.1, 105.7, 107.7, 116.6, 123.3, 124.4, 138.3, 153.8, 174.8; UV (ACN:H_2O 1:3) λ_{max} 217.9, 265.3, 279.6, 289.1 nm; MS-ESI *m/z* 257.1 $[M + Na]^+$, 273.1 $[M + K]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{12}H_{14}N_2O_3Na$ 257.0902, found 257.0903. The ^{1}H NMR matched the literature for 4-methoxy-D-tryptophan in DMSO-d_6, with the exception of the α proton.[115]

4-Methyl-L-tryptophan (25)

4-Methylindole (160 mg, 1.22 mmol) was incubated with 7.5 mg of tryptophan synthase in Buffer TS containing 0.5 mL DMSO ($[DMSO]_{final}$ = 2 %) as described above for 89 h. Column chromatography yielded **25** (175 mg, 0.80 mmol, 66%) as an off-white solid; R_f 0.65 (ACN:H_2O 4:1); ^{1}H NMR (300 MHz, D_2O/NaOD) δ 2.19 (s, 3H), 2.44 (dd, *J* = 8.1, 14.4 Hz, 1H), 2.83 (dd, *J* = 5.4, 14.4 Hz, 1H), 2.96 (dd, *J* = 5.7, 7.8 Hz, 1H), 6.38 (d, *J* = 7.5 Hz, 1H), 6.61 (t, *J* = 7.8 Hz, 1H), 6.69 (s, 1H), 6.86 (d, *J* = 8.1 Hz, 1H); UV (ACN:H_2O 1:3) λ_{max} 219.1, 270.0 nm; MS-ESI *m/z* 219.2 $[M + H]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{12}H_{15}N_2O_2$ 219.1134, found 219.1136. The ^{1}H NMR matched the literature.[100]

4-Vinyl-L-tryptophan (26)

4-vinylindole (**19**, 160 mg, 1.12 mmol) was incubated with 2 mg of tryptophan synthase in a flask equipped with a reflux condenser as described above for 42 h. Column chromatography yielded **26** (180 mg, 0.78 mmol, 70%) as a yellow solid; R_f 0.41 (ACN:H_2O 4:1); ^{1}H NMR (500 MHz, D_2O/NaOD, DSS) δ 2.97 (dd, *J* = 8.0, 14.5 Hz 1H), 3.31 (dd, *J* = 6.0, 14.5 Hz, 1H), 3.53 (dd, *J* = 6.0, 8.0 Hz, 1H), 5.45 (dd, *J* = 1.5, 11.0 Hz, 1H), 5.80 (dd, *J* = 1.5, 17.5 Hz, 1H), 7.20-7.24 (m, 3H), 7.33 (d, *J* = 7.5 Hz, 1H), 7.45 (d,

J = 8.0 Hz, 1H), 7.56 (dd, *J* = 11.0, 17.5 Hz, 1H); ^{13}C NMR (125 MHz, D_2O/NaOD, DSS) δ 35.6, 60.2, 97.5, 114.0, 114.3, 118.5, 119.1, 124.7, 126.5, 128.4, 134.2, 138.1, 139.9, 185.5; UV (ACN:H_2O 1:3) λ_{max} 219.1, 299.8 nm; MS-ESI *m/z* 231.1 $[M + H]^+$, 253.1 $[M + Na]^+$, 269.1 $[M + K]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{13}H_{15}N_2O_2$ 231.1134, found 231.1135.

7-Hydroxy-L-tryptophan (27)

7-Hydroxyindole (160 mg, 1.20 mmol) was incubated with 8 mg of tryptophan synthase as described above for 90 h. Column chromatography yielded **27** (200 mg, 0.91 mmol, 76%) as light gray crystals, which were stored under argon at -80 °C; R_f 0.42 (ACN:H_2O 4:1); ^{1}H NMR (300 MHz, D_2O) δ 3.23 (dd, *J* = 7.8, 15.0 Hz, 1H), 3.63 (dd, *J* = 4.5, 15.0 Hz, 1H), 4.11 (dd, *J* = 4.5, 8.1 Hz, 1H), 6.54 (dd, *J* = 3.9, 4.5 Hz, 1H), 7.05 (s, 1H), 7.07 (d, *J* = 1.2 Hz, 1H), 7.13 (s, 1H); UV (ACN:H_2O 1:3) λ_{max} 219.1, 265.3, 280.7, 290.3 nm; MS-ESI *m/z* 243.1 $[M + Na]^+$; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{11}H_{13}N_2O_3$ 221.0926, found 221.0922. The ^{1}H NMR did not match the literature aromatic region, however it looked similar to 4-hydroxy-L-tryptophan, which would be expected.[114]

7-Methyl-L-tryptophan (28)

7-Methylindole (160 mg, 1.22 mmol) was incubated with 2.34 mg of tryptophan synthase in Buffer TS containing 0.5 mL DMSO ($[DMSO]_{final}$ = 2 %) as described above for 42 h. Column chromatography yielded **28** (230 mg, 1.05 mmol, 86%) as a white solid; R_f 0.47 (ACN:H_2O 4:1); ^{1}H NMR (500 MHz, D_2O) δ 2.57 (s, 3H), 3.36 (dd, *J* = 7.5, 15.0 Hz, 1H), 3.53 (d, *J* = 15.0 Hz, 1H), 4.09-4.15 (m, 1H), 7.15-7.22 (m, 2H), 7.40

(s, 1H), 7.64 (d, J = 7.5 Hz, 1H); UV (ACN:H_2O 1:3) λ_{max} 217.9, 270.0, 276.0 nm; HRMS-ESI TOF m/z $[M + H]^+$ calcd for $C_{12}H_{14}N_2O_2Na$ 241.0953, found 241.0955. The ^1H NMR matched the literature.[100]

Dihydroiso-L-tryptophan (29)

Indoline (160 mg, 1.34 mmol) was incubated in the dark with 1 mg of tryptophan synthase in Buffer TS containing 0.5 mL DMSO ($[DMSO]_{final}$ = 2 %) as described above for 21 h. Column chromatography yielded **29** (240 mg, 1.16 mmol, 87%) as a cream colored solid; R_f 0.57 (ACN:H_2O 4:1); ^1H NMR (500 MHz, D_2O) δ 3.04 (t, J = 8.5 Hz, 2H), 3.39 (dd, J = 8.0, 16.5 Hz, 1H), 3.48 (dd, J = 8.0, 16.5 Hz, 1H), 3.58 (dd, 5.0, 14.5 Hz, 1H), 3.63 (dd, J =7.5, 15.0 Hz, 1H), 4.07 (dd, J = 4.5, 7.0 Hz, 1H), 6.74 (d, J = 8.0 Hz, 1H), 6.86 (t, J = 7.5 Hz, 1H), 7.22 (t, J = 7.5 Hz, 1H), 7.27 (d, J = 7.5 Hz, 1H); ^{13}C NMR (125 MHz, D_2O) δ 28.4, 51.3, 54.0, 54.6, 107.9, 119.3, 125.2, 127.7, 130.9, 152.3. UV (ACN:H_2O 1:3) λ_{max} 205.0, 247.5, 293.8 nm; HRMS-ESI TOF m/z $[M + H]^+$ calcd for $C_{11}H_{15}N_2O_2$ 207.1134, found 207.1148. The ^1H NMR matched the literature, except a contamination singlet at 2.8 ppm; no ^{13}C NMR data was previously reported.[116]

Results and Discussion

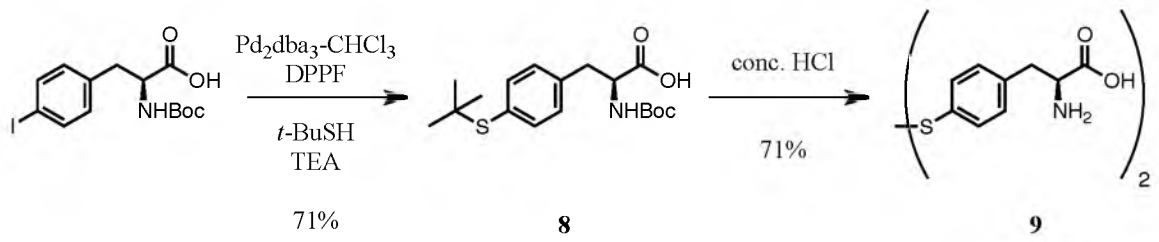
Synthesis of Tyrosine Analogues

O-Dimethylallyl-L-tyrosine (**7**) was synthesized according to the procedure outlined by Pedras and coworkers for use as an authentic standard.[68] Briefly, commercially available *N*-Boc-L-tyrosine was prenylated with dimethylallyl bromide (**4**). Simple amine deprotection by heating at 150 °C and purification by silica gel chromatography yielded pure **7**.

We began the synthesis of 4-mercapto-L-phenylalanine following the literature procedures for the chlorosulfonation of L-phenylalanine followed by a Sn/HCl reduction first reported by Escher and coworkers.[90,91] Although we attempted the above synthesis numerous times, we were never able to obtain the tyrosine analogue. For our purposes, this synthesis also suffered from its lack of isolation and purification of the unprotected amino acid. Both groups used the crude reaction mixture in a variety of protection reactions for further use in peptide synthesis. After searching the literature and testing many other thiol syntheses,[117-125] we settled on a procedure that introduces a *t*-butyl protected thiol[107] that can be deblocked under acidic conditions (Scheme 2.1).[126]

The two-step synthesis of disulfide **9** involves a palladium catalyzed cross-coupling reaction between the commercially available *N*-*t*-Boc-4-iodo-L-phenylalanine and *t*-butylthiol followed by the simultaneous deprotection of the Boc and *S*-*t*-butyl protecting groups with concentrated HCl. Due to the oxidative nature of the deprotection reaction, we isolated the symmetrical disulfide of 4-mercapto-L-phenylalanine (**9**). This was an acceptable product as the disulfide was easily reduced with a reducing agent such as β-mercaptoethanol before addition to an enzymatic assay.

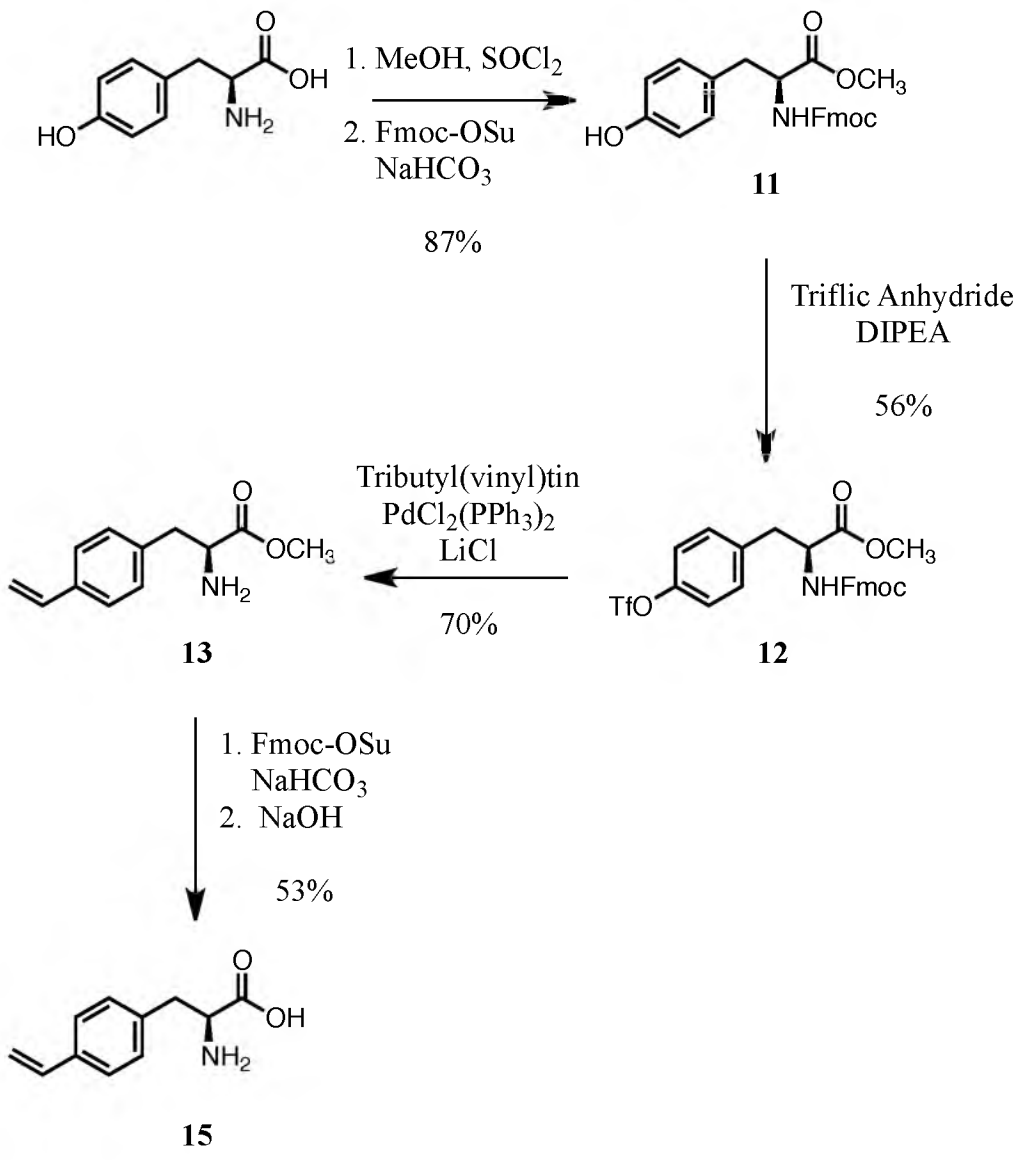
4-Vinyl-L-phenylalanine (**15**) can be synthesized from L-tyrosine by converting the phenolic moiety of tyrosine into a triflic ester which can then be substituted to a vinyl group through the Stille coupling.[92] Due to the fact L-tyrosine *t*-butyl ester is commercially available, we decided to protect the amino acid moiety of L-tyrosine as Boc and *t*-butyl ester moieties, respectively, which allowed both protecting groups to be removed under the same conditions. After protection with Boc, triflic anhydride was used to convert the phenolic alcohol to a triflic ester[127] followed by a Stille coupling with



Scheme 2.1

tributyl(vinyl)tin[92] to obtain *N*-*t*-Boc-4-vinyl-L-phenylalanine *t*-butyl ester. Both the Boc and *t*-butyl ester protecting groups were removed in the presence of trifluoroacetic acid (TFA), but an unwanted side reaction also occurred. NMR analysis of the crude product showed that the vinyl peaks had disappeared. A literature search revealed that styrene polymerizes to polystyrene in the presence of TFA,[128] suggesting that the vinyl amino acid (a styrene derivative) had undergone polymerization.

In order to circumvent this issue, **15** was synthesized using protecting groups that can be removed under nonacidic conditions (Scheme 2.2). *N*-Fmoc-L-tyrosine methyl ester (**11**) was synthesized from L-tyrosine using literature methods and the phenol was converted to its triflate (**12**) under the same procedure as earlier described. Similar to above, the triflate **12** was then treated with tributyl(vinyl)tin to obtain the corresponding vinyl substituted L-tyrosine derivative (**13**). Unfortunately, this reaction also deprotected the Fmoc protecting group to give the free amine, most likely due to diethylamine contamination in the DMF used as a solvent. The free amine was re-protected using Fmoc chemistry to form **14** to facilitate purification by silica gel chromatography. After purification, both the Fmoc and methyl ester were deprotected with 1 N NaOH to obtain 4-vinyl-L-phenylalanine (**15**).

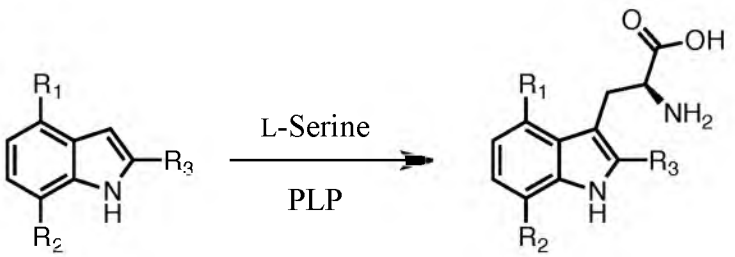

1. MeOH, SOCl2
2. Fmoc-OSu
NaHCO3
87%
11
Triflic Anhydride
DIPEA
56%
12
Tributyl(vinyl)tin
PdCl2(PPh3)2
LiCl
70%
13
1. Fmoc-OSu
NaHCO3
2. NaOH
53%
15

Scheme 2.2

Synthesis of Tryptophan Analogues

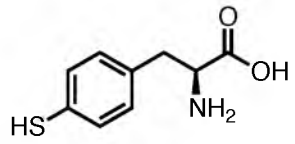
Using the described synthetic pathways for 4-vinyl-L-phenylalanine (**15**) and 4-mercapto-L-phenylalanine as models (see above), we synthesized 4-vinylindole (**19**) and 4-mercaptoindole (**18**) following similar procedures. First, commercially available 4-hydroxyindole was converted to its triflate (**16**).[129] For 4-vinylindole (**19**), triflate ester **16** was converted to a vinyl group by a Stille coupling. For 4-mercaptoindole (**18**), a palladium catalyzed cross-coupling reaction with triflate **16** and *t*-butylthiol followed by acidic deprotection yielded **18**. Both indole analogues were then used to biosynthetically produce tryptophan analogues (see below).

Tryptophan synthase was used to synthesize enantiomerically pure L-tryptophan analogues from a variety of indoles in one simple and convenient step (Scheme 2.3).[100,101] Using both synthetic and commercially available indole analogues, we synthesized the following L-tryptophan analogues using the β-subunit of tryptophan synthase: 2-methyl-tryptophan (**20**), 4-aminotryptophan (**21**), 4-hydroxytryptophan (**22**), 4-mercapto-tryptophan (**23**), 4-methoxytrypotphan (**24**), 4-methyltryptophan (**25**), 4-vinyltryptophan (**26**), 7-hydroxytryptophan (**27**), 7-methyltryptophan (**28**), and dihydroisotryptophan (**29**). The incubations typically gave ~200 mg of purified product. However, if desired, the reactions can be easily scaled up to increase tryptophan production. Other indoles that were not substrates for tryptophan synthase include indene, 4-azaindole (although published),[130] and 4-(*S*-*tert*-butylthio)-indole (**17**). Overall, two L-tyrosine/phenylalanine analogues and ten L-tryptophan analogues were synthesized (Figure 2.1).



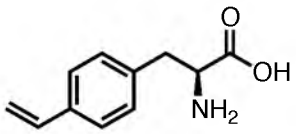
R1: NH_2, OH, SH, OCH_3, CH_3, $CH{=}CH_2$
R2: CH_3, OH
R3: CH_3

Scheme 2.3



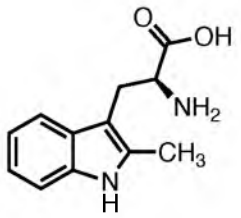
4-Mercapto-L-Phenylalanine

9 (disulfide)



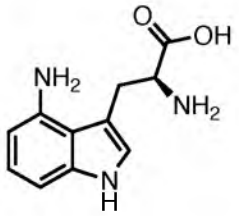
4-Vinyl-L-Phenylalanine

15



2-Methyl-L-Tryptophan

20



4-Amino-L-Tryptophan

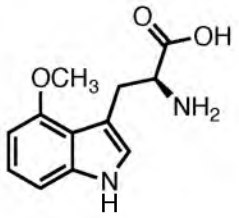
21

4-Hydroxy-L-Tryptophan

22

4-Mercapto-L-Tryptophan

23

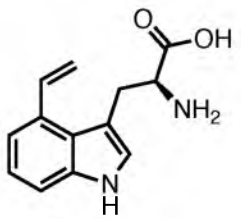


4-Methoxy-L-Tryptophan

24

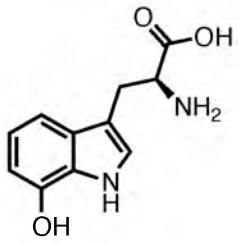
4-Methyl-L-Tryptophan

25



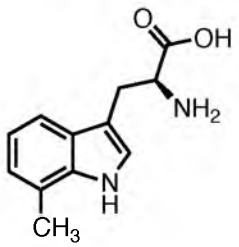
4-Vinyl-L-Tryptophan

26



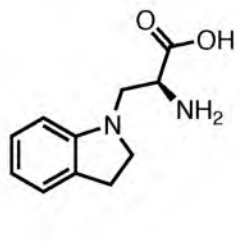
7-Hydroxy-L-Tryptophan

27



7-Methyl-L-Tryptophan

28



Dihydroiso-L-Tryptophan

29

Figure 2.1. Amino acid analogues synthesized.

CHAPTER 3

SIRD: BIOCHEMICAL CHARACTERIZATION AND AROMATIC SUBSTRATE PROMISCUITY

Introduction

Phytopathogenic fungi produce toxic secondary metabolites that provide selective advantages against other microorganisms. Epipolythiodioxopiperazines (ETPs) constitute one class of fungal secondary metabolite toxins. The sulfur-bridged dioxopiperazine ring is constructed from a cyclic dipeptide and is known to contain one, three, or four sulfur atoms.[18] Well-known examples include gliotoxin and sirodesmin PL (Figure 1.1).

The first known genes responsible for the biosynthesis of sirodesmin PL were reported in 2004. Found in the pathogenic organism, *Leptosphaeria maculans*, the sirodesmin PL gene cluster spans a 68 kilobase region encoding enzymes including a two-module nonribosomal peptide synthetase, a thioredoxin reductase, and a prenyltransferase.[64] The prenyltransferase gene, named sirD, was found to be a homologue of a DMATS from *Neotyphodium coenophialum*. Given the other members of the gene cluster and sirodesmin PL's structure, SirD was hypothesized to prenylate the tyrosyl group of either free L-tyrosine or the cyclic dipeptide cyclo-L-Ser-L-Tyr (c-SY) (Scheme 1.3). Pedras and Yu recently completed an isotopic feeding experiment, utilizing tritium-labeled biosynthetic intermediates to determine incorporation into sirodesmin PL. L-Tyrosine and *O*-prenyl-L-tyrosine were efficiently incorporated into phomamide and sirodesmin PL, while was not.[68] This suggested that the SirD-mediated prenylation of L-tyrosine is the first committed step in sirodesmin PL biosynthesis.

A BLAST[69] search reveals that SirD is a member of the aromatic prenyltransferase/DMATS superfamily. 7-DMATS from *A. fumigatus*, 35% identity, is the closest relative to SirD studied to-date. Alignments with other relevant DMATS report the following sequence identities: 31% with FgaPT2 from *A. fumigatus*, 27% with

4-DMATS from *C. purpurea*, and 27% with CTrpPT from *A. oryzae* DSM1147. A putative SirD from *Sirodesmium diversum*, 87% identity, is the nearest known relative of *L. maculans* SirD, however the C-terminus is truncated by 165 residues compared to *L. maculans* SirD.

The mechanism proposed for prenylation by DMATSs is a dissociative electrophilic alkylation of the indole ring by DMAPP where cleavage of the carbon-oxygen bond in DMAPP gives a dimethylallyl cation-PP_i ion pair, with subsequent alkylation of the indole moiety, followed by loss of a proton to give the prenylated product.[80,131] Due to the dissociative electrophilic nature of these enzymes, many aromatic prenyltransferases demonstrate relaxed aromatic substrate specificity.[81,132,133] Of particular interest to this chapter are the indole prenyltransferases 7-DMATS and CTrpPT. 7-DMATS from *A. fumigatus* was cloned, expressed, and reported to prenylate the C7 position of tryptophan's indole ring in a normal orientation.[76] 7-DMATS displayed the ability to prenylate free L- and D-tryptophan (15.5% relative activity to L-tryptophan), as well as the linear (10.9%) and cyclic (1.8%) L-Trp-L-Gly dipeptide. 7-DMATS showed no prenylation product when incubated with L-tyrosine.[76] CTrpPT from *Aspergillus oryzae* DSM1147 was shown to prenylate two different sites on the indole ring of certain L-tryptophan-containing cyclic dipeptides.[77] This enzyme, which shares an identity of 28% with 7-DMATS, catalyzes either normal prenylation at C7 or reverse prenylation at N1. CTrpPT is the first enzyme known to simultaneously prenylate both the C7 and N1 positions of an indole ring.

Given the promiscuous nature of aromatic prenyltransferases, we hypothesized SirD may also accept and prenylate tyrosine, phenylalanine and tryptophan derivatives.

In this chapter, we report the cloning, expression, purification and biochemical characterization of SirD, including the determination of its natural substrates and its ability to prenylate a variety of substituted aromatic amino acids. During the course of our investigations, two studies were published reporting a few of the experiments and results reported in this chapter.[134,135] Where relevant, comparisons and contrasts between our data and the published results will be addressed.

Experimental Procedures

General

NMR experiments were recorded on an INOVA 600 NMR spectrometer equipped with a HCN cryogenic probe and spectra were processed with MestReNova 7.1. All chemical shifts and coupling constants are reported in parts per million (ppm) and hertz (Hz), respectively. ^{1}H chemical shifts are referenced relative to the following peaks: DMSO-d_6 at 2.50 ppm (center peak) or DSS at 0.00 ppm. ^{13}C chemical shifts are referenced relative to the following peaks: DMSO-d_6 at 39.51 (center peak) or DSS at 0.00 ppm. ^{15}N chemical shifts are referenced relative to liquid ammonia in D_2O at δ 0.0 ppm. HPLC was conducted on a Waters 2690 Separation Module equipped with a Microsorb MV^{TM} C18 5 μm column and a Waters 996 photodiode array detector. Normal phase thin layer chromatography (TLC) was carried out on silica gel 60 glass-backed plates (Merck). Reverse phase TLC was carried out on either C8 glass-backed plates (Silicycle) or C18 glass-backed plates with a concentrating zone (Merck). Mass spectrometry was performed by Dr. James Muller at the Chemistry Department Mass Spectrometry Laboratory, University of Utah.

Materials

All compounds were obtained from Sigma-Aldrich unless otherwise noted. The following compounds were purchased from the companies indicated: L-tyrosine [^{14}C(U)], DMAPP triammonium salt [1-^{14}C], IPP triammonium salt [1-^{14}C] (American Radiolabeled Chemicals, Inc.); and ampicillin (Roche). Deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. *E. coli* IDI-1 was provided by Dr. Jonathan Johnston. 2-Oxo-L-tryptophan was supplied by Dr. Raj Viswanathan.

Leptosphaeria maculans and *Sirodesmium diversum* sirD Gene

Design and Cloning

The *L. maculans* sirD gene (NCBI GenBank: AY553235.1) was optimized for heterologous expression in *E. coli* and synthesized by Genscript. The gene (Figure 3.1) was engineered with 5'-BamHI and 3'-HindIII endonuclease restriction sites, an *N*-terminal histidine tag, and a thrombin cleavage site and placed in the pUC57 vector. The modified *L. maculans* sirD gene in pUC57 and the destination vector pET21a(+) were digested with BamHI and HindIII endonucleases (New England Biolabs, Inc.) as per the manufacturer's instructions. After purification by agarose gel electrophoresis, the sirD gene was ligated with T4 DNA Ligase (NEB) into the cut pET-21a(+) vector to produce

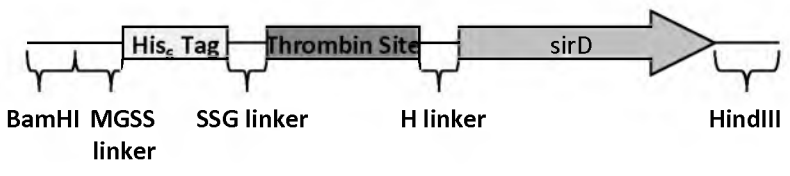


Figure 3.1. The engineered sirD gene constructs synthesized by Genscript.

pLmSirD. pLmSirD was transformed into *E. coli* DH5α (Invitrogen) for plasmid isolation and DNA sequencing and into *E. coli* BL21 (DE3) (Stratagene) for expression. Both *E. coli* strains were grown in liquid LB broth or on solid LB-agar medium at 37 °C. Ampicillin (50 μg mL^{-1}) was used to select vector-containing *E. coli* DH5α and BL21 (DE3) strains.

A putative *S. diversum* sirD gene (AY571972.1) was designed, optimized, synthesized and cloned as described above to produce pSdSirD.

Production and Purification of *L. maculans* SirD

The pLmSirD-BL21 (DE3) cultures were cultivated in 1 L of LB broth containing ampicillin (50 μg mL^{-1}) at 37 °C with shaking at 225 rpm until an OD_{600} of 0.6 was reached. The culture was induced with IPTG to a final concentration of 0.6 mM. After 5 h, cells were harvested by centrifugation (4000 x *g*, 20 min, 4 °C). Pelleted cells were resuspended in lysis buffer (50 mM NaH_2PO_4, pH 8.0, containing 300 mM NaCl and 10 mM imidazole) containing 1 mg mL^{-1} lysozyme and 10 μg mL^{-1} RNase A. After incubation at 0 °C for 30 min, the cell suspension was sonicated for 6 x 10 s at 4 °C. The lysate was clarified by centrifugation (10,000 x *g*, 20 min, 4 °C). The recombinant protein was purified by nickel affinity chromatography (GE Healthcare HisTrap HP) following the manufacturer's instructions. *L. maculans* SirD was eluted with a gradient of 0 – 100% of lysis buffer/elution buffer (50 mM NaH_2PO_4, pH 8.0, containing 300 mM NaCl and 500 mM imidazole). The purified protein was dialyzed 2 x against 4 L of dialysis buffer (20 mM Tris-HCl, pH 8.0, containing 1 mM βME), then 1 x against 4 L of dialysis buffer containing 20% glycerol. The protein concentration was measured using

the Pierce® BCA Protein Assay Kit (Thermo Scientific), divided into 100 µL portions, flash frozen with liquid nitrogen, and stored at -80 °C until use.

Production and Purification of *S. diversum* SirD

The pSdSirD-BL21 (DE3) cultures were cultivated and expressed as described above, but the expressed protein was insoluble. Therefore, the following insolubility troubleshooting procedures were tried: 1) protein denaturation/refolding and 2) coexpression with pGroES/EL.

1) After cell lysis, the pelleted insoluble cellular material was resuspended in urea lysis buffer (100 mM NaH_2PO_4, pH 8.0, containing 10 mM TrisHCl and 8 M urea) with shaking at 100 rpm for 30 min at rt. The suspension was centrifuged (10,000 x *g*, 20 min, rt) and the resulting lysate was shaken with 2.75 mL Ni-NTA slurry (Qiagen) for 60 min at rt. The lysate-resin was loaded onto a column, the flow-through was collected, and the resin was washed/eluted successively with 4 mL of urea buffer C (100 mM NaH_2PO_4, pH 6.3, containing 10 mM TrisHCl and 8 M urea), urea buffer D (100 mM NaH_2PO_4, pH 5.9, containing 10 mM TrisHCl and 8 M urea), and urea buffer E (100 mM NaH_2PO_4, pH 4.5, containing 10 mM TrisHCl and 8 M urea). Purified protein was successively dialyzed against 2 L each of urea dialysis buffer A (20 mM Tris-HCl, pH 8.0, containing 1 mM βME and 4 M urea), urea dialysis buffer B (20 mM Tris-HCl, pH 8.0, containing 1 mM βME, and 2 M urea), urea dialysis buffer C (20 mM Tris-HCl, pH 8.0, containing 1 mM βME and 1 M urea), and urea dialysis buffer D (20 mM Tris-HCl, pH 8.0, containing 1 mM βME and 20% glycerol).

2) The plasmids pSdSirD and pGroESL were co-transformed into BL21 Star™ (DE3) cells (Stratagene) following the manufacturer's instructions. Plasmid pGroESL, a

construct containing the *E. coli* groE operon[136] was obtained from Anthony Gatenby. The cultures were cultivated in 1 L of LB broth containing ampicillin (50 μg mL^{-1}) and chloramphenicol (34 μg mL^{-1}) at 37 °C with shaking at 225 rpm until an OD_{600} of 0.6 was reached. Protein expression was induced with IPTG to a final concentration of 0.5 mM and the culture was cultivated for 4 h. The cell lysis and protein purification procedures were the same as described for *L. maculans* SirD.

Crystallization Studies of *L. maculans* SirD

L. maculans SirD was expressed and purified as described above. Purified *L. maculans* SirD was incubated with 6 units (5 μL) of thrombin (Novagen) at rt for 16 h. The reaction mixture was concentrated to 0.5 mL with a 10 kDa MWCO filter and injected onto a gel filtration column packed with HiLoad Superdex 200 (GE Life Sciences). Thrombin-digested *L. maculans* SirD was purified from thrombin, the cleaved *N*-terminal peptide, and other impurities by elution with 50 mM HEPES, pH 7.5, containing 150 mM NaCl and 1 mM βME. Fractions containing SirD were pooled and concentrated to 9.7 mg/mL. $MgCl_2$ was added to a final concentration of 5 mM and the protein was set in crystallization trays. Crystals appeared after 9 days at rt and promising crystals were looped, placed in the corresponding buffer containing 30% glycerol, and flash frozen by plunging into liquid nitrogen. The following crystals/buffer conditions were mounted and irradiated by Mark Mabanglo in Dr. Chris Hill's laboratory at the Department of Biochemistry, University of Utah: JENA D3, PACT H3a, PACT H3b, PACT H4. Additional crystallization conditions were tested (data not shown), however at best, the results were similar.

Computational Structure Studies of *L. maculans* SirD

The secondary structure of *L. maculans* SirD was predicted from the amino acid sequence using the Advanced Protein Secondary Structure Prediction (APSSP) Server.[137] Each residue was assigned as a helix, strand, or coil along with a probability value for correct prediction. Residues predicted to be in a helix or strand of ≤ 2 consecutive residues were ignored. Combining a primary sequence alignment[138] with the secondary structure prediction for SirD and the secondary structure acquired from the crystal structure of FgaPT2 (4-DMATS from *A. fumigatus*),[59] both the amino acid sequences and the secondary structures of *L. maculans* SirD and the known αββα prenyltransferase FgaPT2 were compared.

Using MODELLER 9v7,[139] a homology model of SirD was created based on the FgaPT2 crystal structure. In order to model tyrosine into the active site in place of tryptophan, the following changes were implemented: 1) deletion of the indole ring side chain from the tryptophan substrate in the FgaPT2 crystal structure while leaving the amino acid moiety intact, 2) construction of a phenol ring side chain and attachment to the remaining amino acid moiety, and 3) energy minimization of the SirD homology model with the new tyrosine substrate.

Natural Substrate Studies

The assay for determining the natural substrates of SirD was conducted in 37 mM Tris-HCl buffer, pH 7.4, containing 14.8 mM $MgCl_2$, 1.1 mM DTT, 0.74 mg/mL BSA, 148 mM KCl, 310 μM ^{14}C-IPP (2.4 mCi/mmol), 1.5 mM of the aromatic substrate, 5.5 μM IDI-1, and 340 nM SirD in a total volume of 135 μL. The negative control reactions were incubated in the absence of IDI-1. Each reaction mixture was incubated at 37 °C

for 30 min. After incubation, 4 μL of the assay mixture were spotted and developed by TLC. For normal phase TLC, a solvent system of 1:9 acetic acid/EtOH was used to develop the plate. The developed TLC plates were imaged on a storage phosphor screen (Molecular Dynamics) and scanned by a Typhoon 8600 Variable Mode Imager (GE Healthcare). The data was visualized and processed with ImageQuant 5.2.

Divalent Cation Studies

The assay was conducted in 50 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$, 0.1 mg/mL BSA, 10 mM KCl, 436 μM ^{14}C-IPP (3.4 mCi/mmol), 4 mM L-tyrosine, 7.5 μM IDI-1, and 500 nM SirD in a total volume of 100 μL. The negative control reaction was incubated in the absence of IDI-1. Each reaction mixture was incubated at 37 °C for 30 min. After incubation, a 3 μL portion of assay mixture was spotted and developed by TLC. TLC development and data visualization was performed as described above.

***O*-Dimethylallyl-L-tyrosine Stability Studies**

The HPLC assay used to determine the stability of *O*-dimethylallyl-L-tyrosine toward quenching conditions was performed on a C18 column with gradient elution using 98:2 H_2O/ACN to 90:10 in 5 min, to 50:50 in 20 min, to 2:98 in 5 min. The total elution volume was 30 mL at a flow rate of 1.0 mL/min. The eluant was monitored at 273.6 nm and product ratios were determined by peak integration. Five minutes before sample injection, 50 μL of *O*-dimethylallyl-L-tyrosine (**7**) and L-tyrosine dissolved in H_2O (unknown concentrations) were added to 50 uL of unbuffered H_2O, 1.5 M trichloroacetic

acid (TCA), or MeOH. After 5 min incubation at rt, each "quenched" mixture was injected for analysis by HPLC.

Thermal Stability Studies

The assay for determining thermal stability was conducted in 20 mM Tris-HCl buffer, pH 8.0, containing 1 mM ^{14}C-L-tyrosine (4 mCi/mmol), 1 mM DMAPP, and 300 (100 for 37 °C) nM SirD in a total volume of 50 μL. The negative control was incubated in the absence of SirD for 2.5 h. Each reaction mixture was incubated at rt, 30 or 37 °C prior to initiation by the addition of DMAPP. For 37 °C, the pre-incubations were 0, 20, 40, 60, 90 and 120 min and the reaction time was 30 min. For rt and 30 °C, the pre-incubations were 0, 10, 20, 30, 60 and 90 min and the reaction times were 10 and 15 min for 30 °C and rt, respectively. The assay was quenched by plunging the sample into a boiling water bath for 15 s. A 10 uL portion was spotted on a C18 RP-TLC plate, which was developed with 25:75 H_2O/MeOH. Data visualization and processing was performed as described above.

Promiscuity Studies

The assays for detecting the promiscuity of aromatic substrates were conducted in 50 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$, 0.1 mg/mL BSA, 10 mM KCl, 1 mM ^{14}C-IPP (0.44-1 mCi/mmol), 1-10 mM of aromatic substrate, ~2 μM IDI-1, and 240-713 nM SirD in a total volume of 50 μL. For thioanalogue assays, 20 mM βME was included. Typically, each reaction mixture was incubated at 30 °C for 1 h and 20 h. After incubation, a 10 μL portion of the assay mixtures was spotted and developed by TLC. For normal phase TLC, a solvent system of 1:9 acetic acid/EtOH or 2:8 H_2O/ACN

was used to develop the plate. Reverse phase C18-TLC and data visualization/processing was performed as previously described.

Kinetic Studies

All kinetic assays were performed in a total volume of 100 μL, incubated at 30 °C for 10 min unless otherwise noted, and quenched by heating at 100 °C for 30 s. After centrifugation, a 10 μL portion of the reaction mixture was spotted and developed by RP-TLC. For L-tyrosine/DMAPP assays, RP-C18 plates were used with a solvent system of 25:75 H_2O/MeOH. For assays with 4-mercapto-L-phenylalanine and tryptophan derivatives, RP-C8 plates were used with a solvent system of 4:6 25 mM NH_4HCO_3/MeOH. Visualization/processing of TLC plates was peformed as previously described. Background radioactivity was determined from incubations without the aromatic substrate. Each kinetic assay was performed in triplicate or quadruplicate.

For L-tyrosine, the assays were conducted in 20 mM Tris-HCl buffer, pH 8.0, containing 80 nM SirD, 1 mM DMAPP, 0.1 μCi of ^{14}C-L-tyrosine, and total L-tyrosine concentrations of 0.04, 0.05, 0.07, 0.08, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35 and 0.4 mM. For DMAPP, the assay was conducted in 20 mM Tris-HCl buffer, pH 8.0, containing 50 nM SirD, 0.5 mM L-tyrosine, 0.05 μCi of ^{14}C-L-tyrosine, and DMAPP concentrations of 0, 0.04, 0.05, 0.06, 0.08, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4 and 0.5 mM. For 4-mercapto-L-phenylalanine, the assay was conducted in 100 mM Tris-HCl buffer, pH 8.0, containing 10 mM βME, 2% v/v glycerol, 5 μM SirD, 1 mM DMAPP, 0.1 μCi of ^{14}C-DMAPP, and 4-mercapto-L-phenylalanine concentrations of 0.25, 0.5, 0.75, 1, 2.5, 5, 7.5 and 10 mM with 30 min incubations. For L-tryptophan, the assay was conducted in 100 mM Tris-HCl buffer, pH 8.0, containing 2% v/v glycerol, 5 μM SirD, 1 mM DMAPP, 0.1 μCi of

^{14}C-DMAPP, and tryptophan concentrations of 0.1, 0.2, 0.3, 0.6, and 1 mM with 20 min incubations. For 4-methoxy and 7-methyltryptophan, concentrations were 0.3, 0.6, 1, 2, 3, 4, 6, and 10 mM with 30 min incubations.

Product Studies

Incubations with tyrosine were conducted in 50 mM Tris-HCl buffer, pH 8.0, containing 2 mM DMAPP, 5 mM L-tyrosine and 2.11 μM SirD in a total volume of 3 mL. The glycerol present in the dialysis buffer was removed by repetitive dialysis/centrifugation using a 10 kDa MWCO filter. The reaction mixture was incubated at 30 °C for 21 h. SirD was removed by centrifugation through a 10 kDa MWCO membrane and the products were purified by chromatography on a C18 RP-HPLC column at a flow rate of 1 mL/min with isocratic elution using ACN:H_2O (25/75). Products were detected by monitoring the UV absorbance at 273.6 nm.

Incubations with thiotyrosine were identical to above with the following differences: 5 mM $MgCl_2$, 20 mM βME, 5 mM DMAPP, 10 mM 4-mercapto-L-phenylalanine and 14.5 μM SirD in a total volume of 5 mL. Products were detected by their absorbance at 254 nm.

Incubations with tryptophan derivatives were identical to above with the following differences: 10 mM $MgCl_2$, 2.92 mM DMAPP, 20 mM tryptophan analogue, 5 μM SirD in a total volume of 5 mL. No βME was included as no thiols were needed to be reduced. The reaction mixtures were incubated for 21-42 h and the products were detected at 225 nm, except for 4-vinyltryptophan (**26**), which was detected at 220 nm. Product ratios were calculated by integrating the area under each product peak visualized at 220 nm (4-vinyl/4-methoxy) or 225 nm (tryptophan/4-methyl/7-methyl).

Structure of Prenylated Aromatic Products

After purification of the prenylated product(s), each product was lyophilized, dissolved in DMSO-d_6 (D, 99.96%) or D_2O (D, 99.96%) (Cambridge Isotope Laboratories, Inc.) and placed in a Norell 3 mm NMR tube (Sigma). Unless noted, the following NMR experiments were performed: 1D ^1H, COSY ^1H-^1H, TOCSY ^1H-^1H, ROESY ^1H-^1H, HSQC or HMQC ^1H-^{13}C, HMBC ^1H-^{13}C, and HMBC ^1H-^{15}N. All spectra were processed with MestReNova 7.1.

***O*-Dimethylallyl-L-tyrosine (30)**

^1H NMR (500 MHz, D_2O/NaOD, DSS) δ 1.74 (s, 3H), 1.78 (s, 3H), 2.77 (dd, J = 7.0, 13.5 Hz, 1H), 2.92 (dd, J = 5.5, 14.0 Hz, 1H), 3.44 (dd, J = 5.5, 7.5 Hz, 1H), 4.59 (d, J = 7.0 Hz, 2H), 5.50 (t, J = 7.0 Hz, 1H), 6.96 (d, J = 8.5 Hz, 2H), 7.19 (d, J = 8.5 Hz, 2H); ^1H NMR (500 MHz, DMSO-d_6) δ 1.69 (s, 3H), 1.73 (s, 3H), 2.77 (dd, J = 8.0, 14.5 Hz, 1H), 3.05 (dd, J = 4.0, 14.5 Hz, 1H), 4.48 (d, J = 6.5 Hz, 2H), 5.41 (t, J = 7.0 Hz, 1H), 6.83 (d, J = 8.5 Hz, 2H), 7.15 (d, J = 8.0 Hz, 2H); UV (ACN:H_2O 1:3) λ_{max} 223.8, 273.6 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{14}H_{19}NO_3Na$ 272.1263, found 272.1257.

***S*-Dimethylallyl-4-thio-L-phenylalanine (31)**

^1H NMR (600 MHz, DMSO-d_6) δ 1.58 (s, 3H), 1.66 (s, 3H), 2.58-2.68 (m, 1H), 2.99 (d, J = 11.4 Hz, 1H), 3.24 (m, 2H), 3.55 (d, J = 7.8 Hz, 2H), 5.24 (t, J = 7.2 Hz, 1H), 7.14 (d, J = 7.2 Hz, 2H), 7.19 (d, J = 7.8 Hz, 2H); ^{13}C NMR (150 MHz, DMSO-d_6) δ 17.3, 25.1, 30.9, 32.1, 56.7, 119.2, 128.2, 129.5, 133.1, 135.5, 137.4; UV (ACN:H_2O 1:3)

λ_{max} 257.0 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{14}H_{19}NO_2NaS$ 288.1034, found 288.1041.

_N_1-(3-Methylbut-1-en-3-yl)-L-tryptophan (32, Reverse _N_-Di-methylallyl-L-tryptophan)

^1H NMR (600 MHz, D_2O) δ 1.62 (s, 6H), 3.06 (dd, *J* = 7.8, 15.0 Hz, 1H), 3.24 (dd, *J* = 4.8, 15.0 Hz, 1H), 3.75 (t, *J* = 6.6 Hz, 1H), 5.02 (d, *J* = 17.4 Hz, 1H), 5.11 (d, *J* = 10.8 Hz, 1H), 6.03 (dd, *J* = 10.8, 17.4 Hz, 1H), 7.05 (t, *J* = 7.2 Hz, 1H), 7.09 (t, *J* = 7.8 Hz, 1H), 7.33 (s, 1H), 7.54 (d, *J* = 8.4 Hz, 1H), 7.60 (d, *J* = 7.8 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 29.6, 30.3, 57.9, 128.5; UV (ACN:H_2O 1:3) λ_{max} 222.7, 284.3 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{16}H_{20}N_2O_2Na$ 295.1422, found 295.1428.

7-(2-Methylbut-2-en-4-yl)-L-tryptophan (33, Normal 7-Di-methylallyl-L-tryptophan)

^1H NMR (600 MHz, D_2O) δ 1.61 (s, 3H), 1.64 (s, 3H), 3.10 (dd, *J* = 7.8, 15.0 Hz, 1H), 3.29 (dd, *J* = 4.8, 15.6 Hz, 1H), 3.47 (d, *J* = 7.2 Hz, 2H), 3.83 (dd, *J* = 4.8, 7.8 Hz, 1H), 5.36 (d, *J* = 7.8 Hz, 1H), 6.98 (d, *J* = 6.6 Hz, 1H), 7.01 (t, *J* = 7.2 Hz, 1H), 7.16 (s, 1H), 7.45 (d, *J* = 7.2 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.7, 27.4, 29.7, 31.8, 57.9, 111.3, 118.9, 122.3, 123.6, 123.8, 127.1, 128.1, 129.5, 137.8, 138.3; UV (ACN:H_2O 1:3) λ_{max} 221.5, 277.2, 287.9 nm; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{16}H_{21}N_2O_2$ 273.1603, found 273.1609.

_N_1-(3-Methylbut-1-en-3-yl)-4-methyl-L-tryptophan (34, Reverse _N_-Dimethylallyl-4-methyl-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.60 (s, 3H), 1.61 (s, 3H), 2.58 (s, 3H), 3.03 (dd, *J* = 9.0, 15.6 Hz, 1H), 3.48 (dd, *J* = 5.4, 15.6 Hz, 1H), 3.68 (dd, *J* = 5.4, 9.0 Hz, 1H), 5.00 (d, *J* = 17.4 Hz, 1H), 5.09 (d, *J* = 10.8 Hz, 1H), 6.01 (dd, *J* = 10.8, 18.0 Hz, 1H), 6.79 (d, *J* = 7.2 Hz, 1H), 6.96 (t, *J* = 7.2 Hz, 1H), 7.30 (s, 1H), 7.38 (d, *J* = 8.4 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 22.2, 29.8, 32.3, 59.6, 61.8, 110.7, 115.2, 116.1, 123.5, 123.8, 128.8, 129.3, 133.7, 138.3, 146.6, 179.5; ^{15}N NMR (60.8 MHz, D_2O) δ 27.9, 145.9; UV (ACN:H_2O 1:3) λ_{max} 225.0, 285.5, 295.0 nm; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{17}H_{23}N_2O_2$ 287.1760, found 287.1762.

4-Methyl-7-(2-methylbut-2-en-4-yl)-L-tryptophan (35, Normal 7-Dimethylallyl-4-methyl-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.61 (s, 3H), 1.63 (s, 3H), 2.56 (s, 3H), 3.06 (dd, *J* = 10.8, 16.2 Hz, 1H), 3.47-3.51 (m, 2H), 3.53-3.59 (m, 1H), 3.74-3.79 (m, 1H), 5.33 (d, *J* = 6.0 Hz, 1H), 6.75 (d, *J* = 7.2 Hz, 1H), 6.86 (t, *J* = 7.2 Hz, 1H), 7.14 (s, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.5, 21.7, 27.3, 31.4, 59.1, 94.6, 123.8, 123.9, 125.6, 127.4, 127.5, 131.0, 137.7, 138.3; UV (ACN:H_2O 1:3) λ_{max} 222.7, 270.0, 290.3, 295.0 nm; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{17}H_{23}N_2O_2$ 287.1760, found 287.1763.

4-Methoxy-7-(3-methylbut-1-en-3-yl)-L-tryptophan (36, Reverse 7-Dimethylallyl-4-methoxy-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.35 (s, 6H), 3.02 (dd, *J* = 8.4, 15.0 Hz, 1H), 3.48 (dd, *J* = 3.6, 15.0 Hz, 1H), 3.84 (s, 3H), 3.88 (dd, *J* = 3.6, 7.8 Hz, 1H), 4.98 (d, *J* = 17.4

Hz, 1H), 5.03 (d, *J* = 10.8 Hz, 1H), 6.00 (dd, *J* = 10.8, 17.4 Hz, 1H), 6.55 (d, *J* = 8.4 Hz, 1H), 6.97 (s, 1H), 7.06 (d, *J* = 7.8 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 29.4, 33.5, 42.0, 57.8, 102.2, 114.4, 119.6, 126.1, 127.8, 149.8; UV (ACN:H_2O 1:3) λ_{max} 220.3, 266.5, 280.7, 291.5 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{17}H_{22}N_2O_3Na$ 325.1528, found 325.1532.

4-Methoxy-7-(2-methylbut-2-en-4-yl)-L-tryptophan (37, Normal 7-Dimethylallyl-4-methoxy-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.59 (s, 3H), 1.60 (s, 3H), 3.02 (dd, *J* =8.4, 15.0 Hz, 1H), 3.35-3.37 (m, 2H), 3.42 (dd, *J* = 4.2, 15.0 Hz, 1H), 3.80 (s, 3H), 3.85 (dd, *J* = 4.8, 8.4 Hz, 1H), 5.30 (t, *J* = 6.6 Hz, 1H), 6.47 (d, *J* = 7.8 Hz, 1H), 6.84 (t, *J* = 7.8 Hz, 1H), 7.00 (s, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.7, 27.5, 31.2, 31.5, 57.8, 59.2, 102.6, 119.3, 121.5, 124.0, 126.5, 137.5, 139.1, 154.7; UV (ACN:H_2O 1:3) λ_{max} 220.3, 266.5, 281.9, 291.5 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{17}H_{22}N_2O_3Na$ 325.1528, found 325.1532.

***N*1-(3-Methylbut-1-en-3-yl)-7-methyl-L-tryptophan (38, Reverse *N*-Dimethylallyl-7-methyl-L-tryptophan)**

^{1}H NMR (600 MHz, D_2O) δ 1.64 (s, 6H), 2.54 (s, 3H), 2.97-3.04 (m, 1H), 3.17-3.23 (m, 1H), 3.66-3.72 (m, 1H), 4.49 (d, *J* = 17.4 Hz, 1H), 4.94 (d, *J* = 10.8 Hz, 1H), 6.30 (dd, *J* = 10.8, 17.4 Hz, 1H), 6.96 (d, *J* = 6.6 Hz, 1H), 7.00 (t, *J* = 7.8 Hz, 1H), 7.43 (s, 1H), 7.46 (d, *J* = 7.8 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 26.6, 34.3, 35.1, 58.2, 62.5, 109.5, 114.7, 119.1, 122.5, 125.5, 128.5, 130.3, 133.3, 137.2, 150.0; UV (ACN:H_2O

1:3) λ_{max} 225.0, 285.5 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{17}H_{22}N_2O_2Na$ 309.1579, found 309.1579.

7-Methyl-6-(2-methylbut-2-en-4-yl)-L-tryptophan (39, Normal 6-Dimethylallyl-7-methyl-L-tryptophan)

^1H NMR (600 MHz, D_2O) δ 1.69 (s, 3H), 1.77 (s, 3H), 2.42 (s, 3H), 3.16 (dd, *J* = 7.8, 15.0 Hz, 1H), 3.36 (dd, *J* = 3.6, 15.0 Hz, 1H), 3.46 (d, *J* = 7.2 Hz, 2H), 3.84 (dd, *J* = 3.6, 12.0 Hz, 1H), 5.32 (t, *J* = 7.2 Hz, 1H), 7.03 (d, *J* = 7.8 Hz, 1H), 7.23 (s, 1H), 7.48 (d, *J* = 7.8 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 14.7, 19.7, 27.5, 33.8, 57.9, 118.4, 121.6, 125.8, 135.8; UV (ACN:H_2O 1:3) λ_{max} 222.7, 271.2, 278.4, 295.0 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{17}H_{22}N_2O_2Na$ 309.1579, found 309.1576.

Timecourse Studies

The assay for time-dependent product formation was conducted in 50 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$, 2.92 mM DMAPP, 20 mM of tryptophan analogue, and 4.7 μM SirD in a total volume of 4 mL. The reaction mixture was incubated at 30 °C. At 1, 2, 3, 5, 10, 15, 30, 60, 120, 240, 360, 480 and 1440 min, 200 μL samples were taken and the enzyme was removed by centrifugation through a 10 kDa MWCO membrane. The filtrates were analyzed by C18 RP-HPLC as described above for product isolation. Products were detected and quantified at 225 nm.

Rearrangement Studies

The assay for the detection of prenyl rearrangement was conducted in 50 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$, 4.7 μM SirD, and an unknown concentration of **34** and **35** in a total volume of 0.5 mL. Additional assays included 2.92

mM DMAPP or 2.92 mM PP_i. The control reaction was incubated in the absence of SirD. The reaction mixture was incubated at 30 °C for 18 h. After protein removal by 10 kDa MWCO filtering, the filtrates were analyzed by C18 RP-HPLC as described above. Products were detected and quantified at 225 nm.

Results and Discussion

pLmSirD and pSdSirD Plasmid Engineering, Expression and Purification

Two plasmids were constructed for overexpression of SirD (Genscript). pLmSirD contained a sirD gene from *Leptosphaeria maculans* optimized for *E. coli* expression (Figure 3.2), while pSdSirD contained an optimized sirD gene from *Sirodesmium diversum* (Figure 3.3). The fidelity of both genes were confirmed by DNA sequencing (University of Utah DNA Sequencing Core Facility) and both genes encoded the NCBI reported protein sequences for *L. maculans* SirD (AAS92554.1) and *S. diversum* SirD (AAT69743.1).

L. maculans SirD and *S. diversum* SirD were overexpressed in *E. coli* BL21 (DE3). SDS-PAGE of *L. maculans* SirD gave a strong band for a soluble protein at ~55 kDa (Figure 3.4), in agreement with the calculated molecular weight of 54,501.8 Da. An electrospray mass spectrum of *L. maculans* SirD gave a molecular weight of 54611.4 Da. SDS-PAGE of the pellet from the induced *S. diversum* SirD culture showed a band at ~38 kDa (data not shown). While in agreement with the calculated molecular weight of 35,617.4 Da, *S. diversum* SirD was insoluble. Two procedures were used in an attempt to solubilize the *S. diversum* SirD protein. First, after cell lysis, the insoluble cellular material was resuspended in a buffer containing 8 M urea, to unfold and solubilize the

```
   1          M  G  S  S  H  H  H  H  H  H  S  S  G  L  V  P  R  G
   1  GGATCCATGGGCAGCAGCCATCATCATCATCATAGCAGCGGCCTGGTGCCGCGTGGC
  21   S  H  M  Q  T  A  R  L  F  Q  G  N  L  N  L  A  A  A  N  I
  61  AGCCATATGCAGACCGCGCGTCTGTTTCAGGGCAACCTGAACCTGGCGGCGGCGAACATT
  41   R  D  Y  E  K  K  E  Q  R  N  L  G  V  W  L  S  L  N  Q  W
 121  CGTGATTATGAGAAGAAAGAACAGCGTAACCTGGGCGTGTGGCTGAGCCTGAACCAGTGG
  61   L  R  L  Y  D  E  D  T  R  F  W  W  T  T  T  A  P  M  L  G
 181  CTGCGTCTGTATGATGAAGATACCCGTTTTTGGTGGACCACCACCGCGCCGATGCTGGGC
  81   R  M  M  E  L  I  G  Y  D  Q  D  A  Q  Q  K  H  L  L  F  Y
 241  CGTATGATGGAACTGATTGGCTATGATCAGGATGCGCAGCAGAAACATCTGCTGTTTTAT
 101   Y  I  Y  V  L  P  S  L  G  R  R  P  S  P  E  G  Y  P  T  G
 301  TATATTTATGTGCTGCCGAGCCTGGGCCGTCGTCCGAGCCCGGAAGGTTATCCGACCGGT
 121   W  N  S  F  M  T  D  D  Y  S  P  L  E  L  S  W  D  W  G  V
 361  TGGAACAGCTTCATGACCGATGATTATAGCCCGCTGGAACTGAGCTGGGATTGGGGCGTG
 141   A  E  G  E  S  S  V  R  F  S  I  E  P  I  G  K  Y  A  G  T
 421  GCGGAAGGCGAAAGCAGCGTGCGTTTTAGCATTGAACCGATTGGCAAATATGCGGGCACC
 161   Q  A  D  P  L  N  Q  K  M  V  Y  Q  L  V  D  G  L  R  P  A
 481  CAGGCAGATCCACTGAACCAGAAAATGGTGTATCAGCTGGTTGATGGCCTGCGTCCGGCG
 181   F  H  H  T  L  D  L  T  L  F  D  V  F  S  E  A  L  T  T  S
 541  TTTCATCATACCCTGGATCTGACCCTGTTTGATGTGTTTAGCGAAGCGCTGACCACCAGC
 201   R  E  K  F  G  T  R  K  L  S  L  E  G  R  S  Q  Y  F  V  A
 601  CGTGAAAAATTTGGCACCCGTAAACTGAGCCTGGAAGGCCGTAGCCAGTATTTTGTGGCG
 221   F  D  L  D  V  G  H  P  R  L  K  A  Y  F  M  P  G  L  K  S
 661  TTTGATCTGGATGTGGGCCATCCGCGTCTGAAAGCGTATTTTATGCCGGGCCTGAAAAGC
 241   I  E  S  N  T  P  V  S  E  L  V  V  K  A  M  D  A  C  E  L
 721  ATTGAAAGCAACACCCCTGTGAGCGAACTGGTGGTGAAAGCGATGGATGCGTGCGAACTG
 261   H  F  G  S  L  F  M  Q  A  F  R  R  L  N  S  D  L  E  A  F
 781  CATTTTGGCAGCCTGTTTATGCAGGCGTTTCGTCGTCTGAACAGCGATCTGGAAGCGTTT
 281   S  A  T  S  Y  H  R  P  E  I  E  I  V  G  I  D  C  V  S  P
 841  AGCGCGACCAGCTATCATCGTCCGGAAATTGAAATTGTGGGCATTGATTGCGTGAGCCCG
 301   V  K  S  R  A  K  I  Y  I  R  H  R  G  T  S  F  D  S  V  C
 901  GTGAAAAGCCGTGCGAAAATTTATATTCGTCATCGTGGCACCAGCTTTGATAGCGTGTGC
 321   R  M  L  S  M  G  A  K  A  P  L  D  A  A  S  V  A  S  L  R
 961  CGTATGCTGAGCATGGGTGCAAAAGCACCGCTGGATGCAGCAAGCGTGGCGAGCCTGCGT
 341   E  L  W  A  L  V  L  G  L  P  K  D  F  P  S  D  Q  E  L  P
1021  GAACTGTGGGCACTGGTGCTGGGTCTGCCGAAAGATTTTCCGAGCGATCAGGAACTGCCG
 361   S  V  P  H  R  T  S  G  V  L  Y  Y  F  E  I  K  P  T  S  D
1081  AGCGTGCCGCATCGTACCAGCGGCGTGCTGTATTATTTTGAAATTAAACCGACCAGCGAT
 381   A  I  V  P  K  V  Y  I  P  V  R  H  Y  A  S  N  D  L  S  I
1141  GCGATTGTGCCGAAAGTGTATATTCCGGTGCGTCATTATGCGAGCAACGATCTGAGCATT
 401   A  Q  G  L  A  T  Y  F  E  R  R  G  Q  T  V  A  A  E  N  Y
1201  GCGCAGGGCCTGGCGACCTATTTTGAACGTCGTGGCCAGACCGTGGCGGCGGAAAACTAT
 421   V  D  A  L  S  D  I  F  S  H  R  S  L  D  S  G  L  G  L  H
1261  GTGGATGCGCTGAGCGATATTTTTAGCCATCGTAGCCTGGATAGCGGCCTGGGCCTGCAT
 441   T  Y  I  S  C  T  F  K  K  T  G  L  S  V  T  S  Y  F  N  P
1321  ACCTATATTAGCTGCACCTTTAAGAAAACCGGCCTGAGCGTGACCAGCTATTTTAACCCG
 461   E  I  Y  H  P  N  R  Y  R  Q  -
1381  GAAATTTATCATCCGAACCGTTATCGTCAGTAAAAGCTT
```

Figure 3.2. ***L. maculans*** **sirD engineered gene and encoded amino acid sequence.**

```
  1          M  G  S  S  H  H  H  H  H  H  S  S  G  L  V  P  R  G
  1  GGATCCATGGGCAGCAGCCATCATCATCATCATCATAGCAGCGGCCTGGTGCCGCGTGGC

 21   S  H  M  Q  T  A  R  L  F  Q  G  N  L  N  L  A  A  A  K  P
 61  AGCCATATGCAGACCGCGCGTCTGTTTCAGGGCAACCTGAACCTGGCGGCGGCGAAACCG

 41   R  E  L  E  K  N  E  Q  G  N  L  G  V  W  L  S  L  N  Q  W
121  CGTGAACTGGAAAAGAACGAACAGGGCAACCTGGGCGTGTGGCTGAGCCTGAACCAGTGG

 61   L  R  L  F  D  E  D  T  R  F  W  W  V  T  T  A  P  I  L  G
181  CTGCGTCTGTTTGATGAAGATACCCGTTTTTGGTGGGTGACCACCGCGCCGATTCTGGGC

 81   R  M  M  E  L  V  G  Y  D  Q  D  A  Q  Q  K  H  L  L  F  Y
241  CGTATGATGGAACTGGTGGGCTATGATCAGGATGCGCAGCAGAAACATCTGCTGTTTTAT

101   Y  I  Y  V  L  P  S  L  G  R  R  P  S  P  E  G  Y  P  T  G
301  TATATTTATGTGCTGCCGAGCCTGGGCCGTCGTCCGAGCCCGGAAGGTTATCCGACCGGT

121   W  N  S  F  M  T  D  D  N  S  P  L  E  L  S  W  D  W  G  I
361  TGGAACAGCTTCATGACCGATGATAACAGCCCGCTGGAACTGAGCTGGGATTGGGGCATT

141   A  E  G  E  S  S  V  R  F  S  I  E  P  I  G  K  Y  A  G  T
421  GCGGAAGGCGAAAGCAGCGTGCGTTTTAGCATTGAACCGATTGGCAAATATGCGGGCACC

161   P  A  D  P  L  N  Q  K  M  V  F  Q  L  V  D  G  L  R  P  A
481  CCGGCAGATCCACTGAACCAGAAAATGGTGTTTCAGCTGGTTGATGGCCTGCGTCCGGCG

181   F  H  H  T  L  D  L  T  V  F  D  V  F  S  E  A  L  T  T  S
541  TTTCATCATACCCTGGATCTGACCGTGTTTGATGTGTTTAGCGAAGCGCTGACCACCAGC

201   R  E  K  L  D  T  R  K  I  S  V  E  G  R  S  Q  Y  F  V  A
601  CGTGAAAAACTGGATACCCGTAAAATTAGCGTGGAAGGCCGTAGCCAGTATTTTGTGGCG

221   F  D  L  D  V  G  H  P  R  L  K  A  Y  F  M  P  G  L  K  A
661  TTTGATCTGGATGTGGGCCATCCGCGTCTGAAAGCGTATTTTATGCCGGGCCTGAAAGCG

241   V  E  A  N  A  S  V  F  E  L  V  V  K  A  M  D  S  C  E  V
721  GTGGAAGCGAACGCGAGCGTGTTTGAACTGGTGGTGAAAGCGATGGATAGCTGCGAAGTG

261   H  F  G  P  L  F  M  Q  A  F  R  R  F  N  S  D  L  A  T  F
781  CATTTTGGCCCGCTGTTTATGCAGGCGTTTCGTCGTTTTAACAGCGATCTGGCGACCTTT

281   S  A  T  S  S  D  R  P  E  V  E  I  V  G  I  D  C  V  I  P
841  AGCGCGACCAGCAGCGATCGTCCGGAAGTGGAAATTGTGGGCATTGATTGCGTGATTCCG

301   A  K  S  R  V  -
901  GCGAAAAGCCGTGTGTAAAAGCTT
```

Figure 3.3. ***S. diversum*** **sirD engineered gene and encoded amino acid sequence.**

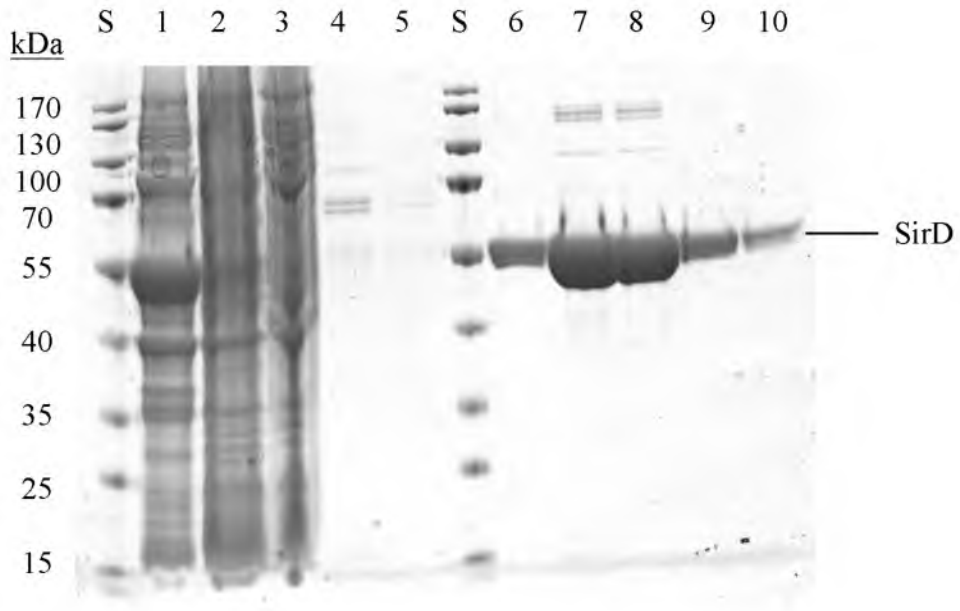


Figure 3.4. SDS-PAGE of pLmSirD expression in BL21 (DE3). Lane S, molecular weight standards; lane 1, pellet; lane 2, cell-free lysate; lane 3, HisTrap column flow-through; lanes 4 and 5, HisTrap column washes; lanes 6-10, HisTrap column elutions.

protein. The soluble portion was then purified under denaturing conditions to obtain purified, unfolded *S. diversum* SirD (Figure 3.5). Refolding was attempted by dialysis against decreasing concentrations of urea. It was not determined whether refolding was successful. Second, *S. diversum* SirD was coexpressed with the GroES/EL chaperone protein complex.[136] SDS-PAGE analysis (data not shown) showed the overexpression of both the *S. diversum* SirD protein and the GroES/EL complex. However, while GroES/EL was found in the soluble lysate, a *S. diversum* SirD band was located in the pellet. A weak band was seen in the soluble washes; however, no corresponding bands were seen in the elution fractions suggesting no His_6 tagged enzyme was present in the soluble lysate. Thus, *S. diversum* SirD was not solubilized.

Due to concerns about *S. diversum*'s folded state, along with the fact that soluble *L. maculans* SirD was in hand, *S. diversum* SirD was set aside. From this point on, *L. maculans* SirD is referred to as simply SirD.

Crystallization Studies

Heterologously expressed SirD was crystallized in a variety of buffer conditions with the help of Mark Mabanglo. SirD most commonly gave rod-shaped crystals, and on occasion gave large diamond-like crystals (Table 3.1). Four crystals were looped, mounted, and irradiated to collect preliminary data sets – three rod-like crystals from PACT H3 and H4 and one large diamond crystal from JENA D3. The three PACT crystals did not diffract. It was initially thought the crystals were not properly aligned in the beam, however, after realignment, the result was unchanged. The JENA D3 crystal gave the diffraction pattern seen in Figure 3.6. A preliminary analysis indicated that the resolution was ~8 Å. Other attempts at crystallization gave thin rods that did not diffract.

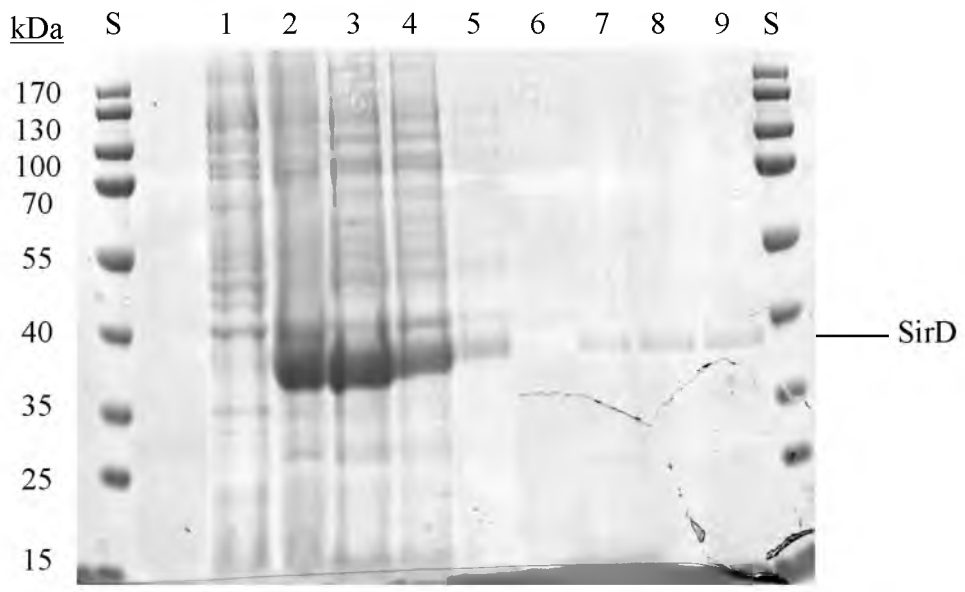


Figure 3.5. SDS-PAGE of pSdSirD expression in BL21 (DE3). Lane S, molecular weight standards; lane 1, cell-free lysate; lane 2, pellet; lane 3, urea lysate; lane 4, HisTrap column flow-through; lanes 5 and 6, HisTrap column washes; lanes 7-9, HisTrap column elutions.

Table 3.1. Crystallization conditions resulting in SirD crystal formation

Well	Crystal Description	Buffer	pH	Salt	PEG
Index H3	rods (ppt.)	0.2 M Na malonate	7.0	-	20% 3350
Index H5	rods (ppt.)	0.1 M succinic acid	7.0	-	15% 3350
Index H7	rods (ppt.)	0.15 M DL-malic acid	7.0	-	20% 3350
PACT G8	irregular	0.1 M BTP	7.5	0.2 M Na sulfate	20% 3350
PACT G9	irregular (ppt.)	0.1 M BTP	7.5	0.2 M Na/K tartrate	20% 3350
PACT G10	long, thin rods	0.1 M BTP	7.5	0.2 M Na/K phosphate	20% 3350
PACT H3	rods	0.1 M BTP	8.5	0.2 M Na iodide	20% 3350
PACT H4	rods	0.1 M BTP	8.5	0.2 M K thiocyanate	20% 3350
JENA B6	irregular	0.1 M Na HEPES	7.5	0.2 M Na acetate	20% 3000
JENA D3	diamond	0.1 M Tris-HCl	8.5	0.5 M LiCl	8% 4000
JENA D9	diamond	0.1 M Na HEPES, 10% 2-propanol	7.5	0.2 M ammonium acetate	16% 400
JENA F7	rods (ppt.)	0.1 M Tris-HCl	8.5	0.5 M LiCl	28% 6000
JENA G2	rods (ppt.)	0.1 M Na HEPES	7.5	0.2 M Ca acetate	10% 8000
HAMPTON H10	rods	0.1 M Tris	7.0	-	20% 2000
HAMPTON H12	rods (ppt.)	0.1 M imidazole	8.0	0.2 M Ca acetate	20% 1000

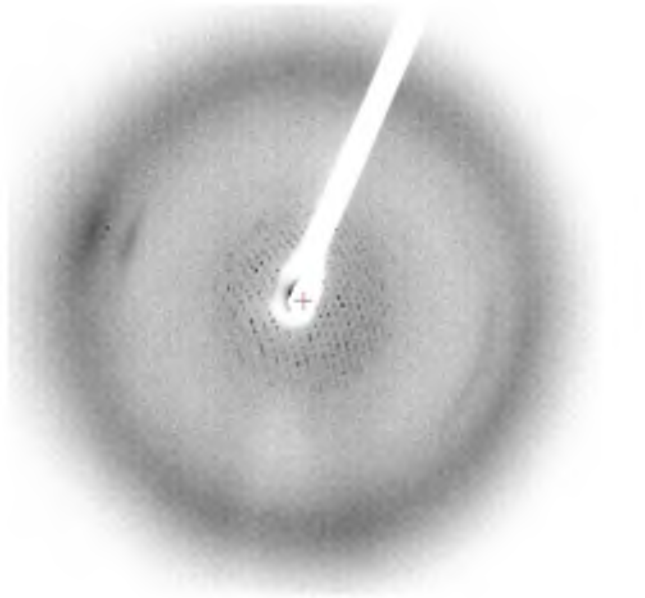


Figure 3.6. Diffraction pattern obtained from the ~8 Å SirD JENA D3 crystal.

Computational Structure Studies

The secondary structure of *L. maculans* SirD was predicted using software on the APSSP Server. Based on the computational prediction, SirD consists of 16 α-helices and13 β-sheets (Figure 3.7). SirD also contains a clear pattern of alternating α-helices and β-sheets. A comparison with the secondary structure of the published crystal structure FgaPT2 shows a strong parallel between the secondary structures of the two aromatic prenyltransferases. Each α-helix or β-sheet in the FgaPT2 structure has an overlapping matching region with the aligned SirD sequence and predicted structure. While there are some minor differences, mainly lengths of helices or sheets, the predicted structure of SirD correlates very well with FgaPT2's known structure suggesting SirD is an αββα prenyltransferase.

In order to visualize SirD, a homology model based on the FgaPT2 crystal structure was constructed (Figure 3.8). The overall structure consists of 12 major α-helices surrounding a 10 antiparallel β-sheet barrel. L-tyrosine and DMASPP are located at the center of the solvent-filled barrel. This tertiary structure correlates well with the predicted secondary structure. Examination of the active site revealed candidates important for binding and/or catalysis (Figure 3.9). Four positively charged residues (R106, K189, R262, K264) are located near the negatively charged diphosphate group in DMASPP. The oxygen atoms of the diphosphate moiety may also be involved in hydrogen bonding with the side chain hydroxyl groups of Y191, Y266, Y330, and Y345, thereby creating additional stabilization of the negative charges. Tyrosine is bound through ionic interactions between R349's side chain and tyrosine's carboxylic acid moiety as well as interactions between the F82 and M83 backbone carbonyls and

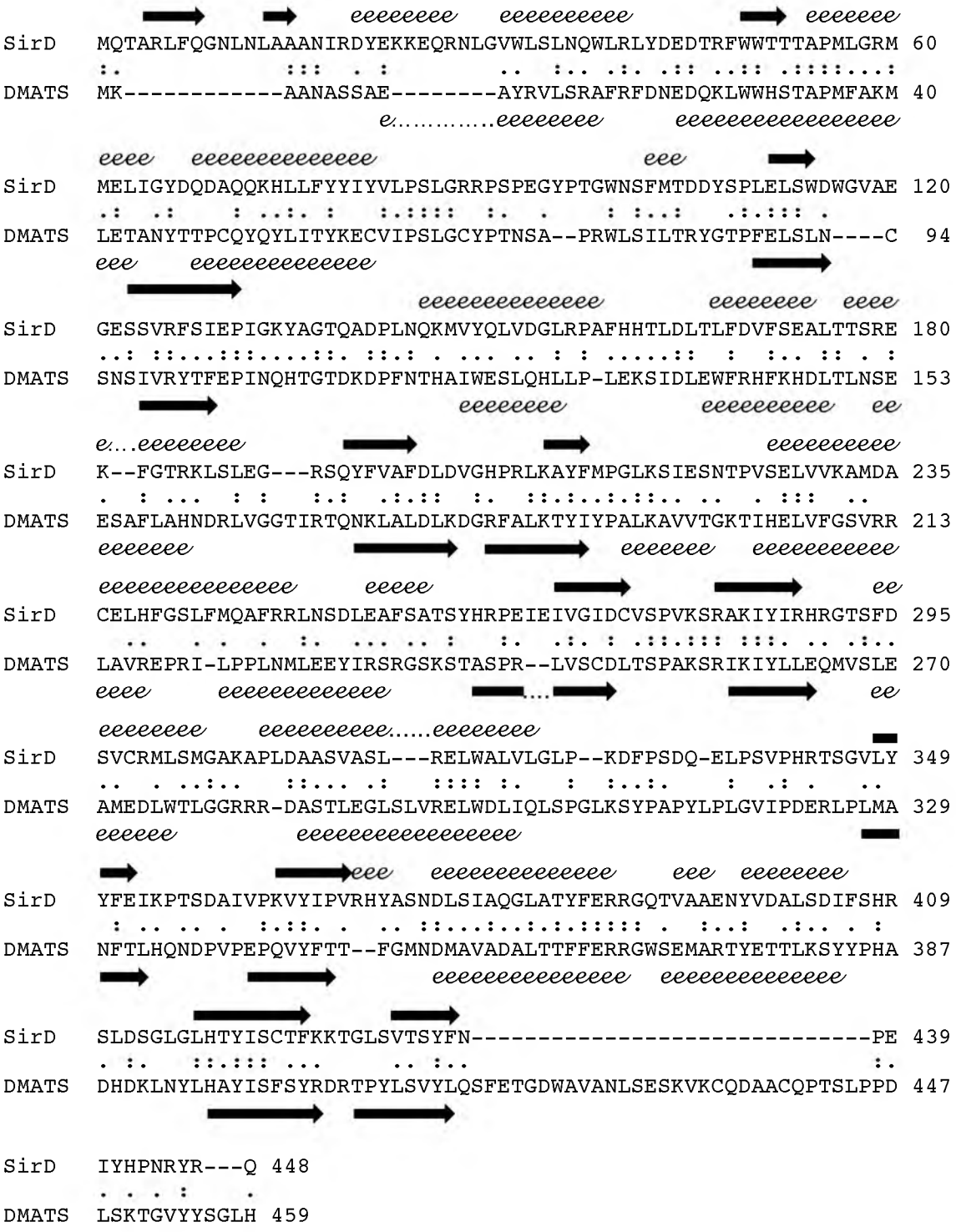


Figure 3.7. A sequence alignment of *L. maculans* SirD and *A. fumigatus* FgaPT2 (4-DMATS) containing the predicted secondary structure of SirD and the secondary structure determined by FgaPT2's crystal structure. Numbers denote the amino acid sequence position; colons (:) and periods (.) signify amino acid sequence identity and similarity, respectively; block arrows denote beta sheets; cursive e's (*e*) denote alpha helices.

Figure 3.8. Homology structure of SirD. (left) Cartoon representation of SirD viewed from the side; (right) cartoon representation of SirD and its two natural substrates L-tyrosine and DMAPP viewed from the top. α-Helices, β-sheets, and flexible loops are shown in cyan, red and magenta, respectively. Bound L-tyrosine and DMASPP are shown in stick representation.

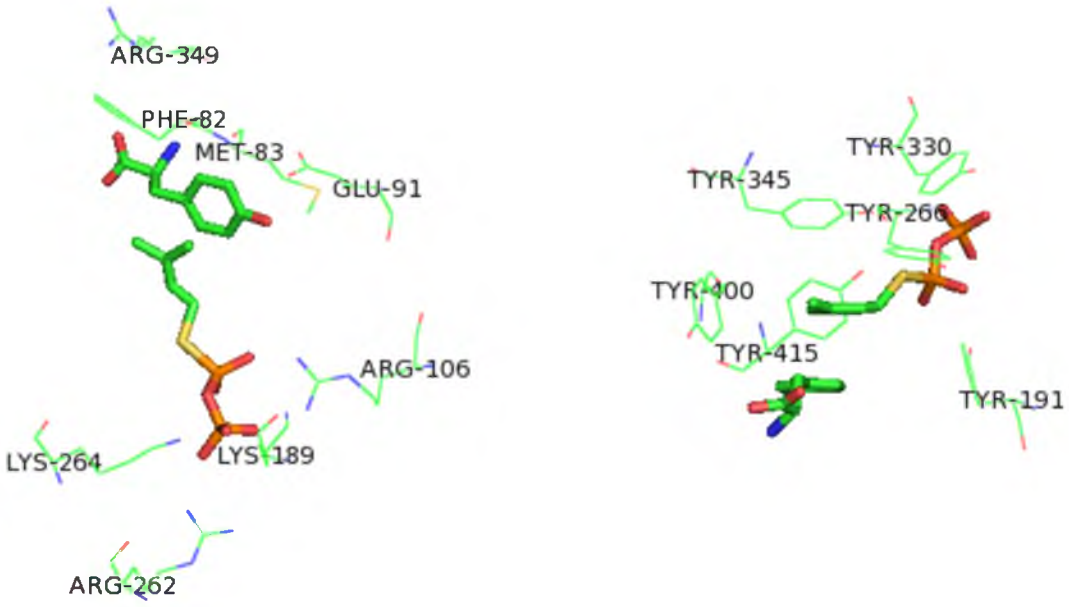


Figure 3.9. Active site of SirD. (left) Active site representation including residues positioned for ionic or hydrogen bond stabilization of the L-tyrosine and DMASPP substrates; (right) active site representation including the tyrosine "shield" positioned around the dimethylallyl moiety of DMAPP. Bound L-tyrosine and DMAPP are shown in stick representation. Active site residues are shown in line representation and labeled with residue name and number.

tyrosine's amino moiety. This orientation positions tyrosine's hydroxyl group in direct line with E91, a potential catalytic base. The distance between the hydroxyl oxygen in tyrosine and C1' of DMASPP is 4.8 Å, suggesting one or both of the substrates must undergo movement to reduce the distance between the atoms involved in the alkylation mechanism. Similar movements were proposed for other aromatic prenyltransferases including FgaPT2[59] and NphB.[56] There is an abundance of tyrosine residues in the active site (Y191, Y266, Y330, Y345, Y400, Y415), many of which surround the allylic substrate. This tyrosine "shield" may protect the reactive carbocation intermediate from interacting with water or other potential reaction-blocking nucleophiles, possibly through π-cation stabilization. Together, these computational structure predictions strongly support SirD's addition into the αββα prenyltransferase family of enzymes.

Natural Substrates and Biochemical Characterization

The natural substrates of SirD were determined using a coupled radio-assay. L-Tyrosine and the cyclic dipeptide c-SY were incubated with ^{14}C-IPP in the presence of IDI-1 to convert ^{14}C-IPP to ^{14}C-DMAPP (Figure 3.10). Negative controls were incubated in the absence of IDI-1, confirming SirD's inability to use IPP as a prenyl donor. The L-tyrosine reaction (lane 5) gave a new spot on the radio-TLC as compared to the L-tyrosine negative control (lane 4), while the c-SY reaction (lane 7) appeared unchanged compared to the c-SY negative control (lane 6). These data clearly indicate that the natural substrates of SirD are L-tyrosine and DMAPP, in agreement with previous biosynthetic pathway labeling experiments.[64,68] An HPLC co-injection of incubation of tyrosine and DMAPP with SirD and *O*-dimethylallyltyrosine (**7**) resulted in a single

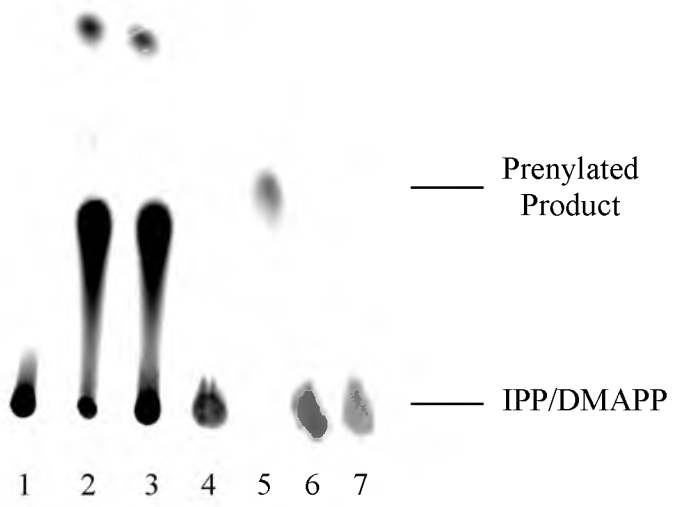
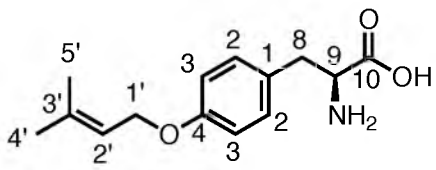
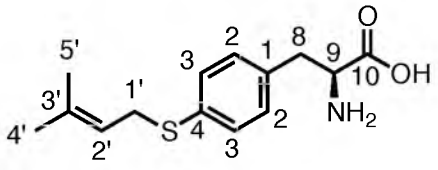


Figure 3.10. Radioautography TLC of the assay for the determination of SirD's natural substrates. Lane 1, ^{14}C-IPP standard; lane 2, ^{14}C-L-tyrosine standard; lane 3, lanes 1 and 2 co-spot; lane 4, L-tyrosine control; lane 5, L-tyrosine reaction; lane 6, cyclo-L-serine-L-tyrosine control; lane 7, cyclo-L-serine-L-tyrosine reaction. Controls and reactions are run in the absence and presence of IDI-1, respectively.

product peak (data not shown), suggesting the enzymatic and synthetic prenylated products were identical.

To confirm *O*-prenylation, L-tyrosine was incubated with DMAPP and SirD and one product, **30**, was formed. The ^{1}H NMR spectrum in D_2O (Table 3.2) confirmed normal prenylation by the presence of proton signals for methylene protons at C1' (4.59 ppm, d, J = 7.0 Hz), the proton at C2' (5.50 ppm, t, J = 7.0 Hz), and methyl groups at 1.74 and 1.78 ppm. An AB quartet for the aromatic protons at 6.96 and 7.19 ppm, each integrating to two protons, indicated that the aromatic ring was *para*-substituted. The downfield chemical shift (4.59 ppm) for the protons at C1' indicates that the dimethylallyl moiety is attached to oxygen. Thus, SirD catalyzes the normal *O*-

Table 3.2. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from L-tyrosine and 4-mercapto-L-phenylalanine

Structure/ Compound	 **30**		 **31**	
Position	**δ_H, multi., *J* (D_2O)**	**δ_H, multi., *J* (DMSO-d_6)**	**δ_H, multi., *J***	**δ_C**
1	-	-	-	137.4
2	7.19, d, 8.5	7.15, d, 8.0	7.14, d, 7.2	129.5
3	6.96, d, 8.5	6.83, d, 8.5	7.19, d, 7.8	128.2
4	-	-	-	133.1
8	2.92, dd, 5.5, 14.0	3.05, dd, 4.0, 14.5	2.99, d, 11.4	32.1
	2.77, dd, 7.0, 13.5	2.77, dd, 8.0, 14.5	2.58-2.68, m	-
9	3.44, dd, 5.5, 7.5	shielded by H_2O	3.24, m	56.7
10	-	-	-	n.f.
1'	4.59, d, 7.0	4.48, d, 6.5	3.55, d, 7.8	30.9
2'	5.50, t, 7.0	5.41, t, 7.0	5.24, t, 7.2	119.2
3'	-	-	-	135.5
4'	1.78, s	1.73, s	1.66, s	25.1
5'	1.74, s	1.69, s	1.58, s	17.3

prenylation of L-tyrosine. Furthermore, HRMS of the enzymatic product shows a peak at *m/z* 272.1257, corresponding to a monoprenylated sodium adduct of tyrosine.

Concomitant with our work, Kremer and coworkers reported that SirD is a tyrosine *O*-prenyltransferase.[134] They established the structure of the product by ^{1}H NMR spectroscopy, however, they used DMSO-d_6 as the solvent instead of D_2O. To verify our product was identical to the reported product, we re-performed NMR analysis using DMSO-d_6 as the solvent. With the exception of the α proton shielded by a moisture peak, the ^{1}H NMR spectrum matched the literature spectrum in DMSO-d_6.

We examined the dependence of the reactions catalyzed by SirD on Mg^{2+}. TLC analysis showed spots for product in both the presence and absence of Mg^{2+} (data not shown). Thus, SirD does not require Mg^{2+}. Other soluble aromatic prenyltransferases are independent of divalent cations.[55,76,78]

The thermal stability of SirD at rt, 30 and 37 °C was determined by pre-incubation followed by an assay for activity (Figure 3.11). As expected, the enzyme was more stable at lower temperatures. After 20 min, SirD retains 90%, 72%, and 54% relative activities at rt, 30 and 37 °C, respectively. After 90 min, the respective activities decrease to 62%, 59%, and 22%.

Stability of *O*-Dimethylallyl-L-tyrosine under Acidic Conditions

The stability of *O*-dimethylallyl-L-tyrosine (**7**) was determined in unbuffered water, water at pH ~2, and in methanol using tyrosine as an internal standard. As a control, the compounds were incubated with unbuffered H_2O and the product ratios were determined to be 34% and 66% for tyrosine and **7**, respectively. There was no change in the ratio of tyrosine and **7** when the compounds were incubated in methanol, indicating

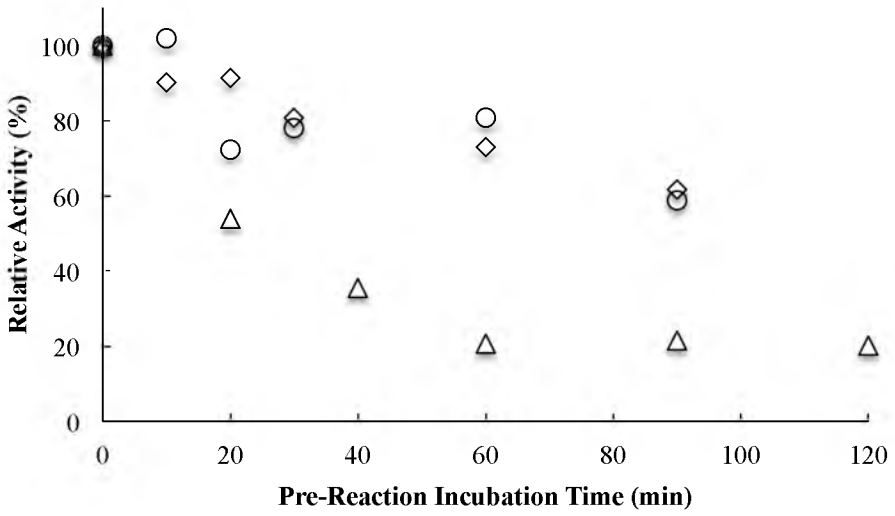


Figure 3.11. Thermal stability of SirD. After SirD incubation at room temperature (◇), 30 °C (○) or 37 °C (△) for designated lengths of time, enzymatic activities were determined relative to the 0 min pre-reaction incubation samples.

that no degradation of **7** had occurred. Conversely, incubation with 1.5 M trichloroacetic acid resulted in 61% and 39% of tyrosine and **7**, respectively. This drastic change in product ratios indicates that **7** is not stable under acidic conditions. Therefore, acidic conditions should not be used, especially to quench enzyme kinetic experiments.

Kinetic Studies with L-Tyrosine and DMAPP

Michaelis-Menten kinetic parameters for SirD's natural substrates L-tyrosine and DMAPP were determined. Rates (k_{cat}) and Michaelis constants (K_M) were determined from a nonlinear regression fit by GraFit 5.0.11. V_{max} = 0.49 μmol min^{-1} mg^{-1}, corresponding to a turnover number of 27 min^{-1}, which compared well to the published value of 0.58 μmol min^{-1} mg^{-1}.[134] K_M = 0.14 mM for L-tyrosine, also compared well with the previously reported values of 0.13 mM[134] and 0.19 mM.[135] K_M = 0.040 mM for DMAPP, ~ 4 times less than the reported value.[134] This difference may arise due to our

exclusion of divalent cations. DMAPP was not tested in the presence of divalent cations, however, evidence of increased K_Ms for DMAPP was reported for 4-DMATS in the presence of divalent cations.[84] Michaelis-Menten plots and kinetic parameters are summarized in Figure 3.12 and Table 3.3, respectively.

Promiscuity Studies

Commercially available analogues of tyrosine and tryptophan were tested as alternate substrates in the presence of DMAPP and SirD by the qualitative radio-TLC coupled assay outlined above. SirD accepted tyrosine and tryptophan analogues (Figure 3.13). SirD prenylated tyrosine, 4-aminophenylalanine, and 3,4-dihydroxyphenylalanine; SirD did not prenylate the *ortho* and *meta* isomers of tyrosine. Thus, a prenyl acceptor *para* (C4) to the amino acid moiety is required for prenylation. SirD did not prenylate phenol, even in the presence of alanine and/or glycine, indicating the importance for the aromatic substrate to contain an amino acid moiety. This is undoubtedly due to the necessary ionic and polar interactions between the active site residues and the substrate's amino acid moiety, thus binding and positioning the substrate in a catalytic manner. Given the sequence similarity between SirD and members of the DMATS superfamily, SirD prenylating tryptophan and tryptophan analogues was expected. Similar rates of turnover for tryptophan and 4-methyltryptophan suggested prenylation did not occur at C4 of the indole ring (similar to 4-DMATS and FgaPT2).

Two studies reported SirD's capacity to accept and prenylate a variety of aromatic substrates including 4-aminophenylalanine, 3,4-dihydroxyphenylalanine, tryptophan and 4-methyltryptophan.[134,135] Using NMR, Li and coworkers identified normal prenylation at the 4-NH_2 and 4-OH positions of the tyrosine derivatives[135] and the C7 position of

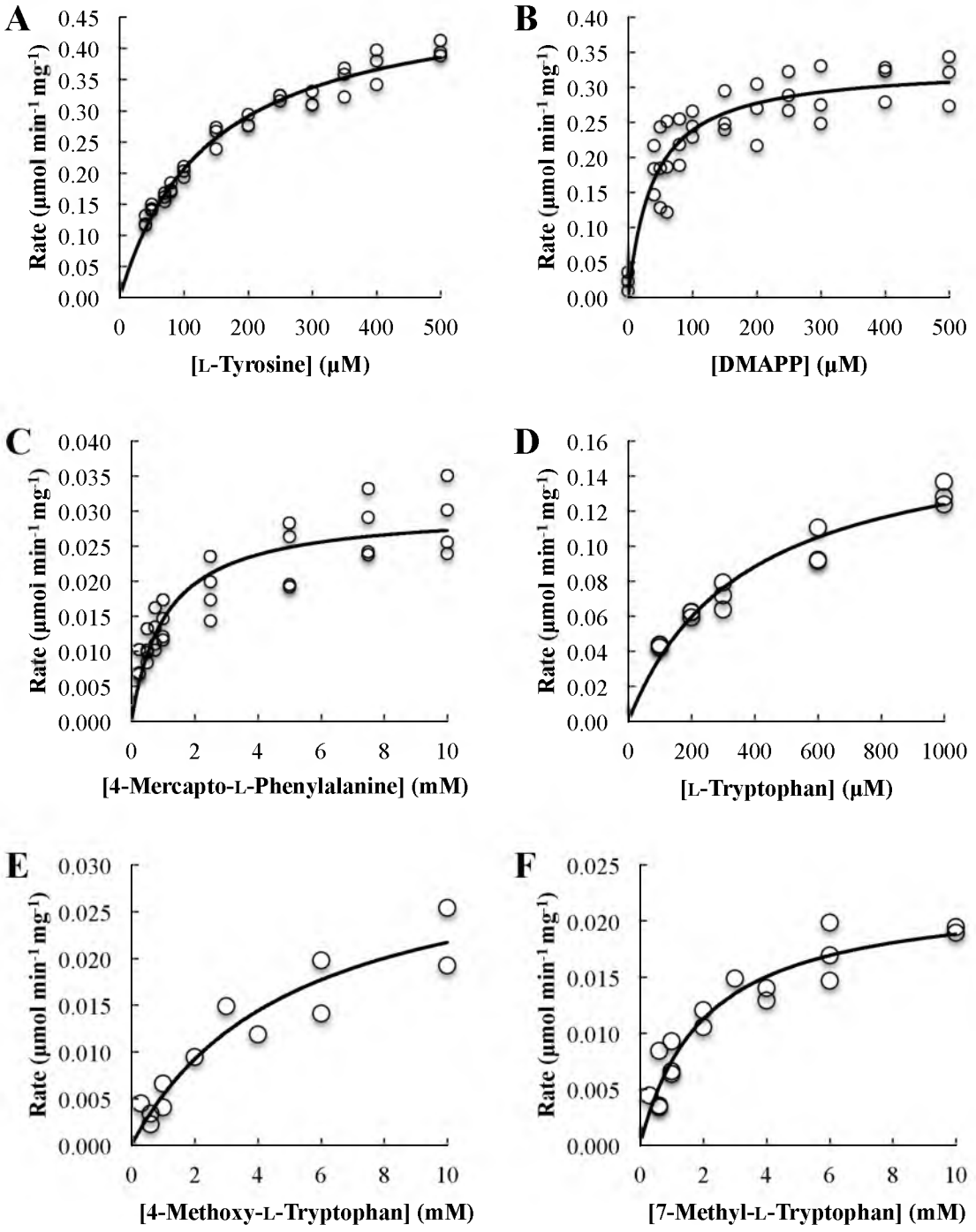


Figure 3.12. Michaelis-Menten kinetic plots for select SirD substrates. Each graph plots substrate concentration vs. rate and illustrates individual data points and a nonlinear regression curve. A, L-tyrosine; B, DMAPP; C, 4-mercapto-L-phenylalanine; D, L-tryptophan; E, 4-methoxy- L-tryptophan; F, 7-methyl- L-tryptophan. All kinetic experiments were conducted at 30 °C.

Table 3.3. Kinetic parameters for SirD substrates at 30 °C

Substrate	V_{max} (μmol min^{-1} mg^{-1})	K_M (mM)	k_{cat} (min^{-1})	k_{cat} / K_M (min^{-1} mM^{-1})
L-TYR	0.49 ± 0.01	0.14 ± 0.01	27 ± 1	190
DMAPP	0.34 ± 0.02	0.04 ± 0.01	18 ± 1	450
L-TYR (repeat)	1.4 ± 0.1	0.30 ± 0.04	76 ± 6	260
4-SH-L-PHE (**9**)	0.030 ± 0.002	1.1 ± 0.2	1.6 ± 0.1	1.5
L-TRP	0.17 ± 0.01	0.37 ± 0.06	5.7 ± 0.1	15
4-OCH_3-L-TRP (**24**)	0.033 ± 0.006	5.1 ± 2.0	0.77 ± 0.20	0.15
7-CH_3-L-TRP (**28**)	0.023 ± 0.002	2.0 ± 0.4	0.77 ± 0.07	0.38

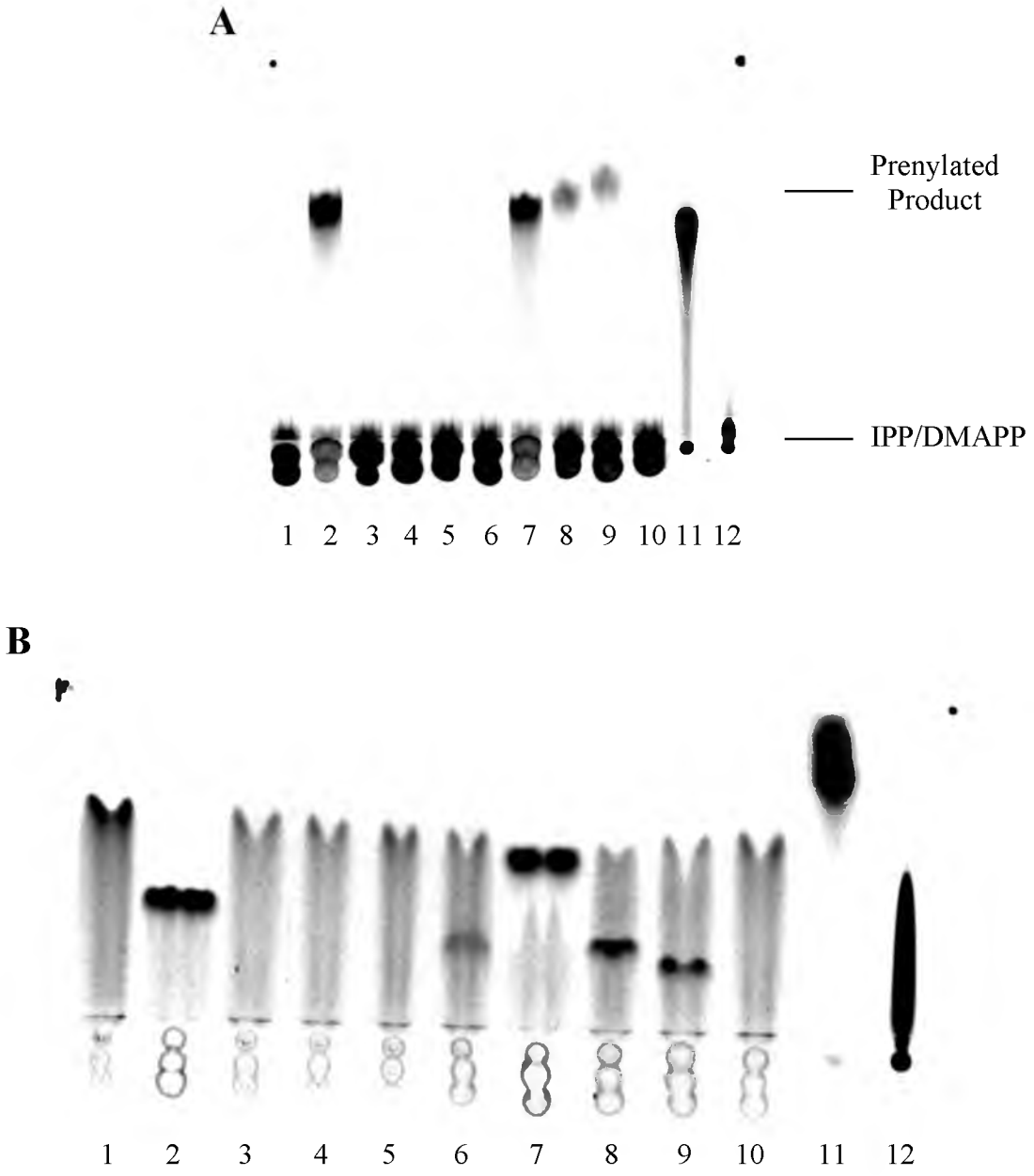


Figure 3.13. Radioautography TLCs (A, normal phase; B, reverse phase) of substrate promiscuity assays for SirD. Lane 1, L-tyrosine negative control; lane 2, L-tyrosine positive control; lane 3, DL-*o*-tyrosine; lane 4, DL-*m*-tyrosine; lane 5, *O*-methyl-L-tyrosine; lane 6, 4-amino-L-phenylalanine; lane 7, 3,4-dihydroxy-L-phenylalanine; lane 8, L-tryptophan; lane 9, 4-methyl-DL-tryptophan; lane 10, phenol + L-glycine + L-alanine; lane 11, ^{14}C-L-tyrosine standard; lane 12, ^{14}C-IPP standard. Each aromatic substrate was tested at a final concentration of 1 mM.

tryptophan.[134] Thus, SirD was shown to prenylate oxygen, nitrogen, and carbon atoms on aromatic amino acids.

The analogues 4-mercapto-L-phenylalanine (**9**) and 4-vinyl-L-phenylalanine (**15**) were also tested as described above. Due to **9**'s intrinsic insolubility in water, **9** was dissolved in DMSO or minimal NaOH prior to its addition to the assay. The appearance of an aromatic product spot after incubations of 1 h and 20 h confirmed the prenylation of 4-mercapto-L-phenylalanine by SirD (Figure 3.14A). All assays containing **9** showed turnover, except the assay containing 20% v/v DMSO (lane 2^c). To our knowledge, this is the first report of an aromatic prenyltransferase performing an *S*-prenylation. Conversely, no products appeared when **15** was incubated with DMAPP and SirD in the coupled assay or analyzed by HPLC (data not shown). The vinyl moiety is either too large for the active site, not properly positioned for prenylation, or the alkene is electronically unable to attack the dimethylallyl cation.

The following tryptophan analogues were tested as described above and found to be substrates of SirD: 4-amino (**21**), 4-hydroxy (**22**), 4-methoxy (**24**), 4-vinyl (**26**), 7-methyl (**28**), and dihydroisotryptophan (DIT, **29**). Some substrates showed turnover at concentrations as low as 1 mM (**26** and **28**) while others only showed turnover at concentrations of 10 mM (**21**, **24**, **29**) (Figure 3.14). Because 7-DMATS is the closest relative to SirD and SirD was reported to prenylate C7 of tryptophan,[134] the products of the 4-substituted analogues were hypothesized to be normal C7 prenylated. The prenylation site of the 7-methyl analogue **28** was of particular interest as the C7 position was blocked, preventing the assumed normal C7 prenylation. Other aromatic analogues determined not to be substrates for SirD include 2-methyltryptophan (**20**), 2-oxo-

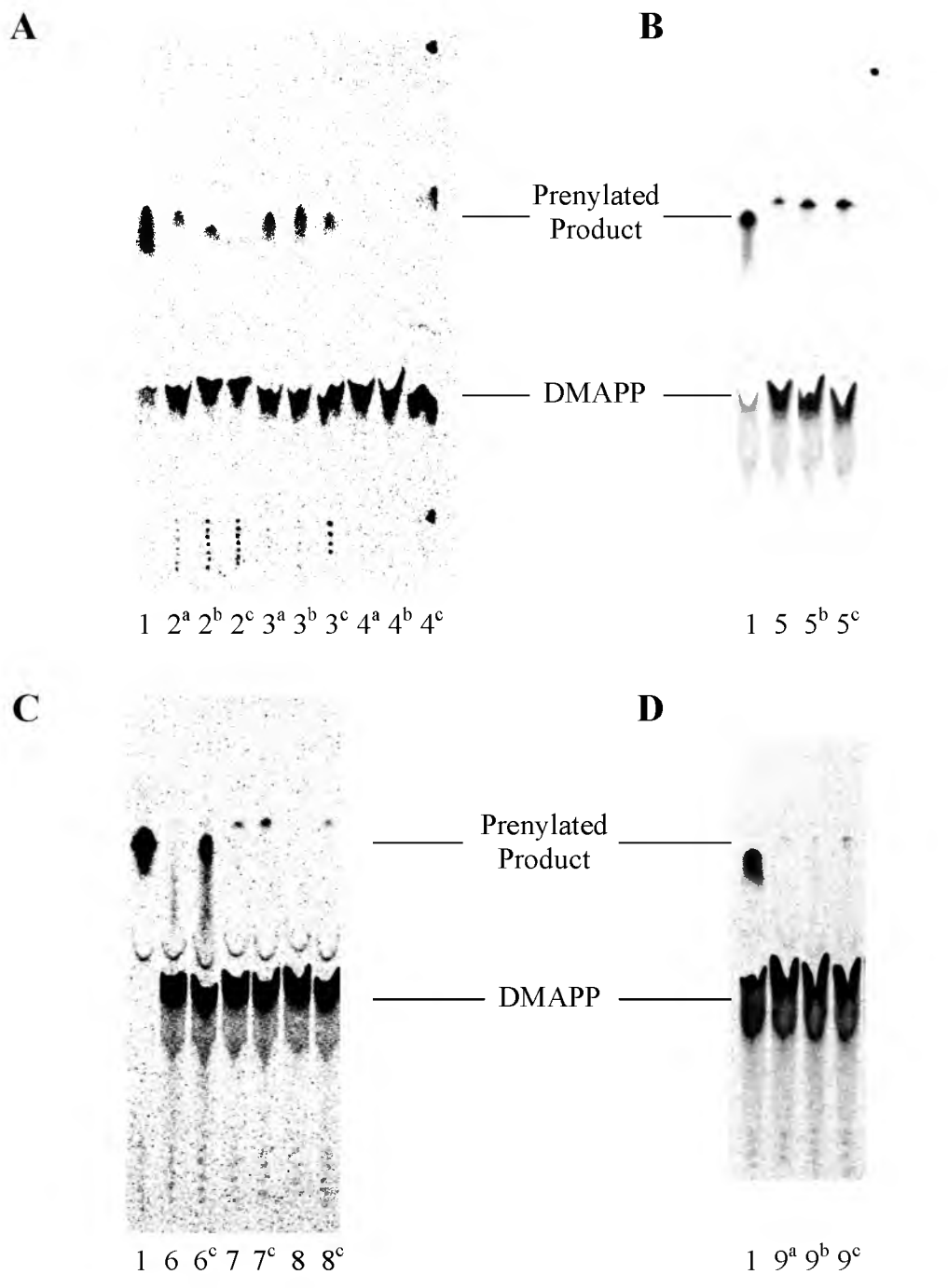


Figure 3.14. Radioautography TLCs of substrate promiscuity assays for SirD. Lanes 1, L-tyrosine positive control; lanes 2, 4-mercapto-L-phenylalanine in DMSO; lanes 3, 4-mercapto-L-phenylalanine in NaOH; lanes 4, 4-methoxy-L-tryptophan; lanes 5, 4-vinyl-L-trytophan; lanes 6, 4-amino-L-tryptophan; lanes 7, 7-methyl-L-tryptophan; lanes 8, dihydroiso-L-tryptophan; lanes 9, 4-hydroxy-L-tryptophan. Superscripts designate aromatic substrate concentrations: a, 2.5 mM; b, 5 mM; c, 10 mM; no superscript, 1 mM.

tryptophan, 4-methylphenylalanine, 7-hydroxytryptophan (**27**), 4- and 7-hydroxyindoles and indene (Table 3.4, TLC data not shown). Unsurprisingly, non-aromatic amino acids including cysteine, serine, and threonine are not substrates for SirD.

Kinetic Studies with Tyrosine and Tryptophan Analogues

Michaelis-Menten kinetic parameters for 4-mercapto-L-phenylalanine, L-tryptophan, 7-methyl-L-tryptophan and 4-methoxy-L-tryptophan were determined (Figure 3.12, Table 3.3). Rates (k_{cat}) and Michaelis constants (K_M) were determined as described above. Previously, we determined that K_M for L-tyrosine was 0.14 mM. After the synthesis of **9**, **24**, and **28**, we repeated the experiments with tyrosine and freshly expressed SirD. K_M = 0.30 mM for L-tyrosine, while K_M = 1.1 mM for thiotyrosine. k_{cat} = 1.6 min^{-1} for thiotyrosine, substantially smaller than that of tyrosine (76 min^{-1}). Thus, the catalytic efficiency (k_{cat}/K_M) for thiotyrosine was less than 1% of the normal reaction.

The increase in K_M and the decrease in k_{cat} of **9** compared with tyrosine indicate that replacing oxygen with sulfur decreases substrate binding and reduces the rate of the reaction. The pKa of thiotyrosine's sulfhydryl provides a reasonable explanation. The pKa of tyrosine's hydroxyl is ~10, causing the phenolic oxygen to be protonated in physiological conditions. As described above, tyrosine's hydroxyl group was modeled to be in direct line with E91, a potential catalytic base. E91 is hypothesized to abstract the hydrogen from tyrosine's hydroxyl, facilitating oxygen attack of the dimethylallyl cation. The pKa of thiotyrosine's sulfhydryl is ~7, potentially causing the sulfur to be deprotonated in physiological conditions. A negatively charged sulfur atom positioned near E91 would cause ionic repulsion, thereby decreasing substrate binding affinity and impeding nucleophilic attack.

Table 3.4. Aromatic substrate analogues tested for turnover with SirD and corresponding product ratios

Compound	[Compound] mM	Reaction (y/n)	Product Ratios (%)
L-Tyrosine*	1	y	100
DL-*o*-Tyrosine*	1	n	-
DL-*m*-Tyrosine*	1	n	-
O-Methyl-L-tyrosine	1	n	-
4-Amino-L-phenylalanine*	1	y	100
3,4-Dihydroxy-L-phenylalanine*	1	y	100
4-Mercapto-L-phenylalanine (**9**) in DMSO	2.5, 5, 10	y, y, n	100
4-Mercapto-L-phenylalanine (**9**) in NaOH	2.5, 5, 10	y, y, y	100
4-Methyl-L-phenylalanine	1	n	-
4-Vinyl-L-phenylalanine (**15**)	5, 10	n, n	-
L-Tryptophan*[a]	1	y	7:93
4-Amino-L-tryptophan (**21**)	1, 10	n, y	n.d.
4-Hydroxy-L-tryptophan (**22**)	2.5, 5, 10	y, y, y	n.d.
4-Mercapto-L-tryptophan (**23**)	1	n	-
4-Methoxy-L-tryptophan[b] (**24**)	2.5, 5, 10	n, n, y	8:92
4-Methyl-L-tryptophan*[a] (**25**)	1	y	79:21
4-Vinyl-L-tryptophan[b] (**26**)	1, 5, 10	y, y, y	52:48
7-Hydroxy-L-tryptophan (**27**)	2.5, 5, 10	n, n, n	-
7-Methyl-L-tryptophan*[a] (**28**)	1,10	y, y	13:87
2-Methyl-L-tryptophan (**20**)	1	n	-
Dihydroiso-L-tryptophan (**29**)	1, 10	n, y	100
2-Oxo-L-tryptophan	1	n	-
4-Hydroxyindole	4	n	-
7-Hydroxyindole	4	n	-
Indene	10	n	-
Phenol	10	n	-
L-Alanine	10	n	-
L-Glycine	10	n	-
Phenol + L-Alanine + L-Glycine	1 (each)	n	-
L-Cysteine	10	n	-
L-Serine	10	n	-
L-Threonine	10	n	-
Cyclo-L-Tyr-L-Ser*	1	n	-

*Aromatic substrates previously reported[134,135]
[a]Product ratios determined by peak integration at 225 nm
[b]Product ratios determined by peak integration at 220 nm
n.d., not determined

K_M^{Trp} = 0.37 mM is comparable with the previously reported value,[134] as well as the value for tyrosine. K_M = 2.0 and 5.1 mM for the 7-methyl and 4-methoxy analogues, respectively, were substantially higher. k_{cat} for the tryptophan analogues ranged between 0.77 and 5.7 min^{-1}, substantially smaller than k_{cat} for tyrosine (76 min^{-1}). The catalytic efficiency (k_{cat}/K_M) for tryptophan was 6% of the normal reaction, comparable with the reported relative yield of 8.1%.[134] The k_{cat}/K_M for the 7-methyl and 4-methoxy analogues were less than 1% of the normal reaction. The saturation limit of 4-methyltryptophan was not found (linear up to 10 mM, data not shown), preventing the determination of k_{cat} and K_M using Michaelis-Menten kinetics. The maximum solubility of 4-methyltryptophan for this study was 10 mM.

Product Studies

4-Mercapto-L-phenylalanine (9)

Given that SirD can catalyze prenylations on oxygens, nitrogens, and carbons, prenylation at the *para*-sulfur atom of **9** by SirD seemed likely. To confirm our hypothesis, a large-scale incubation of **9** and DMAPP with SirD was completed. C18 RP-HPLC was used to identify and purify the aromatic product **31** (10.3 min) (Figure 3.15). Monoprenylation of **9** was confirmed using HRMS – the sodium adduct of dimethylallyl thiotyrosine was found to have an *m/z* of 288.1041. We established the structure of **31** using a combination of ^1H and 2D NMR spectroscopy (Table 3.2). Normal prenylation was confirmed by the presence of signals for the methylene protons at C1' (3.55 ppm, d, J = 7.8 Hz), the proton at C2' (5.24 ppm, t, J = 7.2 Hz), and the methyl groups at C3' (1.66 and 1.58 ppm). An AB quartet for the aromatic protons at 7.14 and 7.19 ppm, each integrating to two protons, indicated that the benzene ring was

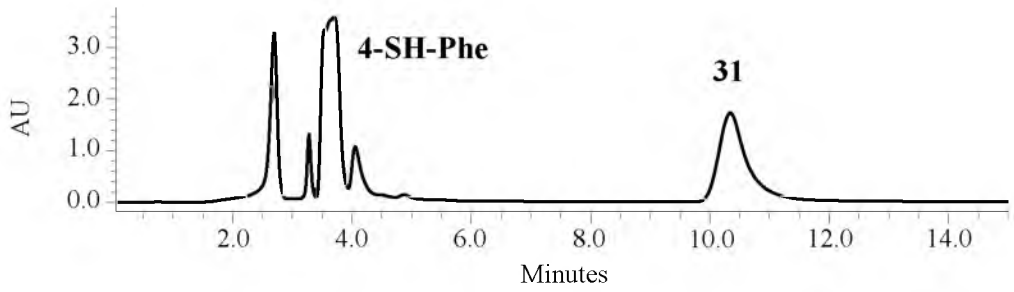


Figure 3.15. HPLC chromatogram of 4-mercapto-L-phenylalanine and DMAPP incubation with SirD. Product formation was monitored at 254 nm.

para-substituted. Thus, aromatic prenylation occurred at the sulfur atom. The HMBC spectrum revealed correlations between the dimethylallyl's methylene protons and C2' (119.2 ppm), C3' (135.5 ppm), and the *para* carbon C4 (133.1 ppm). These data prove that SirD catalyzes a normal *S*-prenylation of 4-mercapto-L-phenylalanine. Therefore, SirD catalyzes normal *O*-, *N*-, and *S*-prenylations of tyrosine derivatives.

4-Methyl-L-tryptophan (25)

SirD catalyzes prenylation on L-tryptophan and 4-methyl-DL-tryptophan. SirD was reported to normal prenylate the C7 position of tryptophan[134] and tryptophan derivatives.[135] The structure of 7-dimethylallyltryptophan was confirmed by ^{1}H NMR spectroscopy,[76,134] while the identity of 7-dimethylallyl-4-methyltryptophan was established by HPLC.[135] Having 4-methyl-L-tryptophan (**25**) in hand and no 7-dimethylallyl-4-methyltryptophan NMR data in the literature, we decided to confirm C7 prenylation of **25** by NMR spectroscopy.

Surprisingly, incubation with **25** and DMAPP with SirD gave two products, **34** (18.8 min, 79%) and **35** (24.6 min, 21%) (Figure 3.16). HRMS analysis revealed that the molecular ions of the compounds were 68 Da higher than **25**, indicating the addition of one isoprene unit.

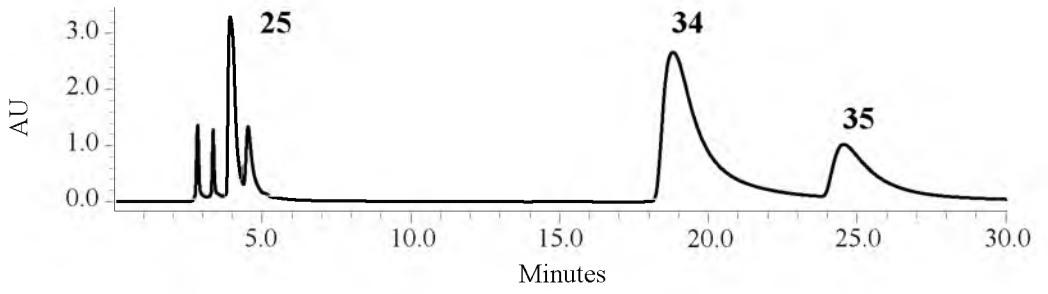


Figure 3.16. HPLC chromatogram of 4-methyl-L-tryptophan and DMAPP incubation with SirD. Product formation was monitored at 225 nm.

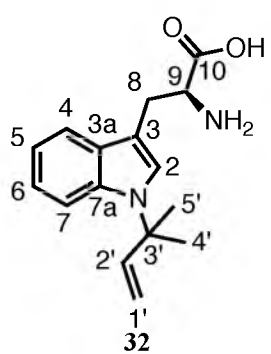
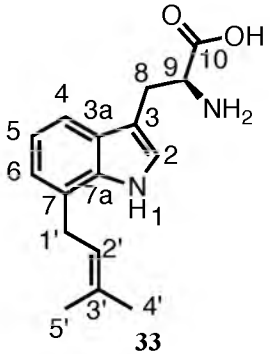
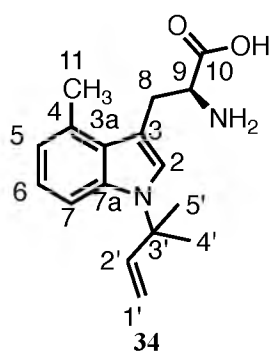
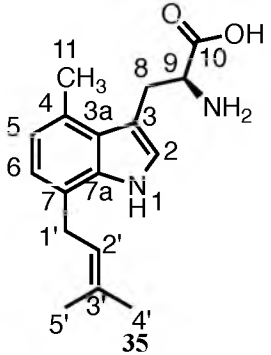
Since 7-dimethylallyl-4-methyltryptophan was reported to be the sole product from incubation of 4-methyltryptophan and DMAPP with SirD,[135] we expected one of the products to be 7-prenylated. NMR analysis (Table 3.5) confirmed **35** was 7- prenylated due to the presence of aromatic proton signals at C2 (7.14 ppm, s), C5 (6.75 ppm, d, J =7.2 Hz), and C6 (6.86 ppm, d, J = 7.2 Hz). Normal prenylation was confirmed by the presence of signals for the methylene protons at C1' (3.47-3.51 ppm, m), the proton at C2 (5.33 ppm, t, J = 6.0 Hz), and the methyl groups at 1.61 and 1.63 ppm. The HMBC spectrum showed correlations between the methylene protons at C1' and the C7 carbon of the indole ring (125.6 ppm). These data verify that SirD catalyzes a normal *C*-prenylation of 4-methyltryptophan at the C7 position of the indole ring.

In contrast to **35**, the ^{1}H NMR spectrum of **34** established that the indole ring contained a reverse dimethylallyl moiety with three characteristic vinyl signals at 6.01 ppm (dd, J = 10.8, 18.0 Hz), 5.09 ppm (d, J = 10.8 Hz) and 5.00 ppm (d, J = 17.4 Hz) for the protons at C2' and C1', respectively, and methyl groups at 1.61 and 1.60 ppm. The ROESY spectrum displayed a correlation between the methyl groups of the prenyl moiety and the proton at C2 (7.30 ppm). The presence of four aromatic protons at C2, C4 (6.79 ppm), C5 (6.96 ppm), and C6 (7.38 ppm) indicated prenylation occurred at N1 of the indole ring. Together, these data prove **34** is reverse *N*1-dimethylallyl-4-methyl-L-tryptophan.

L-Tryptophan

Given the surprising result that SirD catalyzes prenylations at two positions on the indole ring of 4-methyltryptophan, we decided to reinvestigate the reaction with L-tryptophan to verify one prenylation product or establish the presence of two prenylation

Table 3.5. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from L-tryptophan and 4-methyl-L-tryptophan

Structure/ Compound	 **32**		 **33**		 **34**		 **35**	
Position	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**
1	-	-	exc.	-	-	-	exc.	-
2	7.33, s	128.5	7.16, s	127.1	7.30, s	128.8	7.14, s	127.5
3	-	n.f.	-	111.3	-	110.7	-	94.6
3a	-	n.f.	-	129.5	-	129.3	-	127.4
4	7.60, d, 7.8	n.f.	7.45, d, 7.2	118.9	-	133.7	-	131.0
5	7.05, t, 7.2	n.f.	7.01, t, 7.2	122.3	6.79, d, 7.2	123.5	6.75, d, 7.2	123.8
6	7.09, t, 7.8	n.f.	6.98, d, 6.6	123.6	6.96, t, 7.2	123.8	6.86, d, 7.2	123.8
7	7.54, d, 8.4	n.f.	-	128.1	7.38, d, 8.4	115.2	-	125.6
7a	-	n.f.	-	138.3	-	138.3	-	138.3
8	3.24, dd, 4.8, 15.0	30.3	3.29, dd, 4.8, 15.6	29.7	3.48, dd, 5.4, 15.6	32.3	3.53-3.59, m	n.f.
	3.06, dd, 7.8, 15.0	-	3.10, dd, 7.8, 15.0	-	3.03, dd, 9.0, 15.6	-	3.06, dd, 10.8, 16.2	-
9	3.75, t, 6.6	57.9	3.83, dd, 4.8, 7.8	57.9	3.68, dd, 5.4, 9.0	59.6	3.74-3.79, m	59.1
10	-	n.f.	-	n.f.	-	179.5	-	n.f.
11	-	-	-	-	2.58, s	22.2	2.56, s	21.7
1'	5.11, d, 10.8	n.f.	3.47, d, 7.2	31.8	5.09, d, 10.8	116.1	3.47-3.51, m	31.4
	5.02, d, 17.4	-	-	-	5.00, d, 17.4	-	-	-
2'	6.03, dd, 10.8, 17.4	n.f.	5.36, t, 7.8	123.8	6.01, dd, 10.8, 18.0	146.6	5.33, t, 6.0	123.9
3'	-	n.f.	-	137.8	-	61.8	-	137.7
4'	1.62, s	29.6	1.61, s	27.4	1.61, s	29.8	1.61, s	27.3
5'	1.62, s	29.6	1.64, s	19.7	1.60, s	29.8	1.63, s	19.5

δ_H and δ_C in ppm; *J* in Hz; n.f. = not found; exc. = exchanged

products. Incubation of tryptophan and DMAPP with SirD gave products **32** (12.7 min, 7%) and **33** (15.3 min, 93%) (Figure 3.17). HRMS analysis confirmed both compounds were monoprenylated products.

NMR analysis (Table 3.5) revealed that compound **33** had a ^{1}H NMR spectrum very similar to the reported data of 7-dimethylallyltryptophan.[76] The HMBC spectrum showed a correlation between the methylene protons at C1' (3.47 ppm, d, J = 7.2 Hz) and C7 of the indole ring (128.1 ppm). Together, these data verified that **33** is normal 7-dimethylallyltryptophan.

Compound **32** was determined to contain a reverse dimethylallyl moiety as evidenced by signals of the three vinyl protons at 6.03 ppm (dd, J = 10.8, 17.4 Hz), 5.11 ppm (d, J = 10.8 Hz) and 5.02 ppm (d, J = 17.4 Hz), and the methyl groups at 1.62 ppm. Although the NMR sample was too dilute to collect relevant TOCSY, HMQC, or HMBC data, the ROESY spectrum showed a correlation between the proton at C2 (7.33 ppm, s) and the prenyl moiety's methyl groups. The presence of 5 aromatic protons at C2, C4 (7.60 ppm), C5 (7.05 ppm), C6 (7.09 ppm), and C7 (7.54 ppm) confirmed prenylation occurred at a position other than the available indole carbons, namely the indole nitrogen. Collectively, these data prove **32** is formed by reverse prenylation at the N1 position of tryptophan. Therefore, SirD catalyzes normal prenylation at C7 or reverse prenylation at N1 of tryptophan and 4-methyltryptophan.

Simultaneous prenylation was recently seen with CTrpPT, catalyzing both normal C7 and reverse N1 prenylations on tryptophan-containing cyclic dipeptides.[77] An amino acid sequence alignment revealed that SirD shares 27% identity with CTrpPT from *A.*

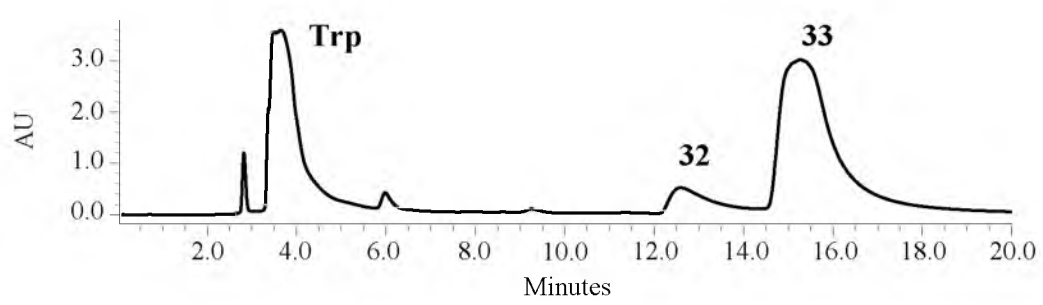


Figure 3.17. HPLC chromatogram of L-tryptophan and DMAPP incubation with SirD. Product formation was monitored at 225 nm.

oryzae DSM1147.[140] This sequence similarity suggests SirD and CTrpPT share important aspects of both their active sites and reaction mechanisms.

One hypothesis for "simultaneous" prenylation is alkylation at one site, N1 or C7, followed by an aza-Claisen rearrangment to form the other product.[89,141,142] To investigate this hypothesis, we performed a 24 h timecourse experiment by incubating L-tryptophan or **25** with DMAPP and SirD. At each time point, a quenched portion of assay mixture was analyzed by HPLC and the ratio of products was determined by integrating the area under each peak. Overall, we saw no evidence for an aza-Claisen rearrangement (Figure 3.18). Significantly, the product ratios were present in expected amounts after only 1 min of incubation. In addition, incubation with **34** or **35** as a substrate did not yield any other products. SirD was unable to rearrange either product alone (Figure 3.19), or in the presence of DMAPP or inorganic pyrophosphate (data not shown). These experiments support a direct electrophilic alkylation mechanism, where the positioning of the indole nucleus relative to the dimethylallyl cation determines the site and orientation of prenylation.

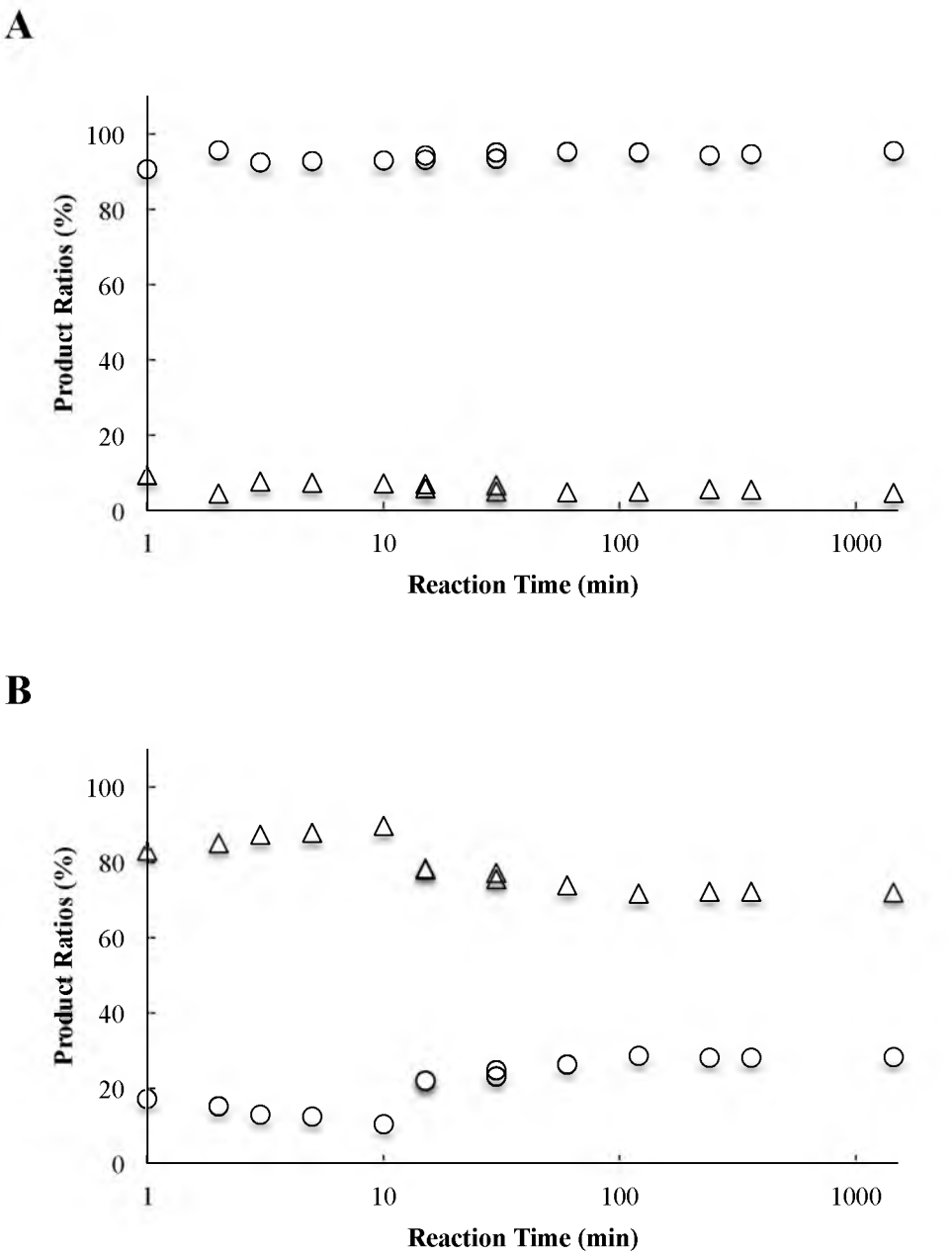


Figure 3.18. Timecourse of L-tryptophan (A) or 4-methyl-L-tryptophan (B) incubations with DMAPP and SirD. A, compounds 32 and 33 are depicted as △ and ○, respectively; B, compounds 34 and 35 are depicted as △ and ○, respectively. Product ratios were determined by peak integration at 225 nm. X-axes were plotted on a $\log_{10}$ scale for data point clarity.

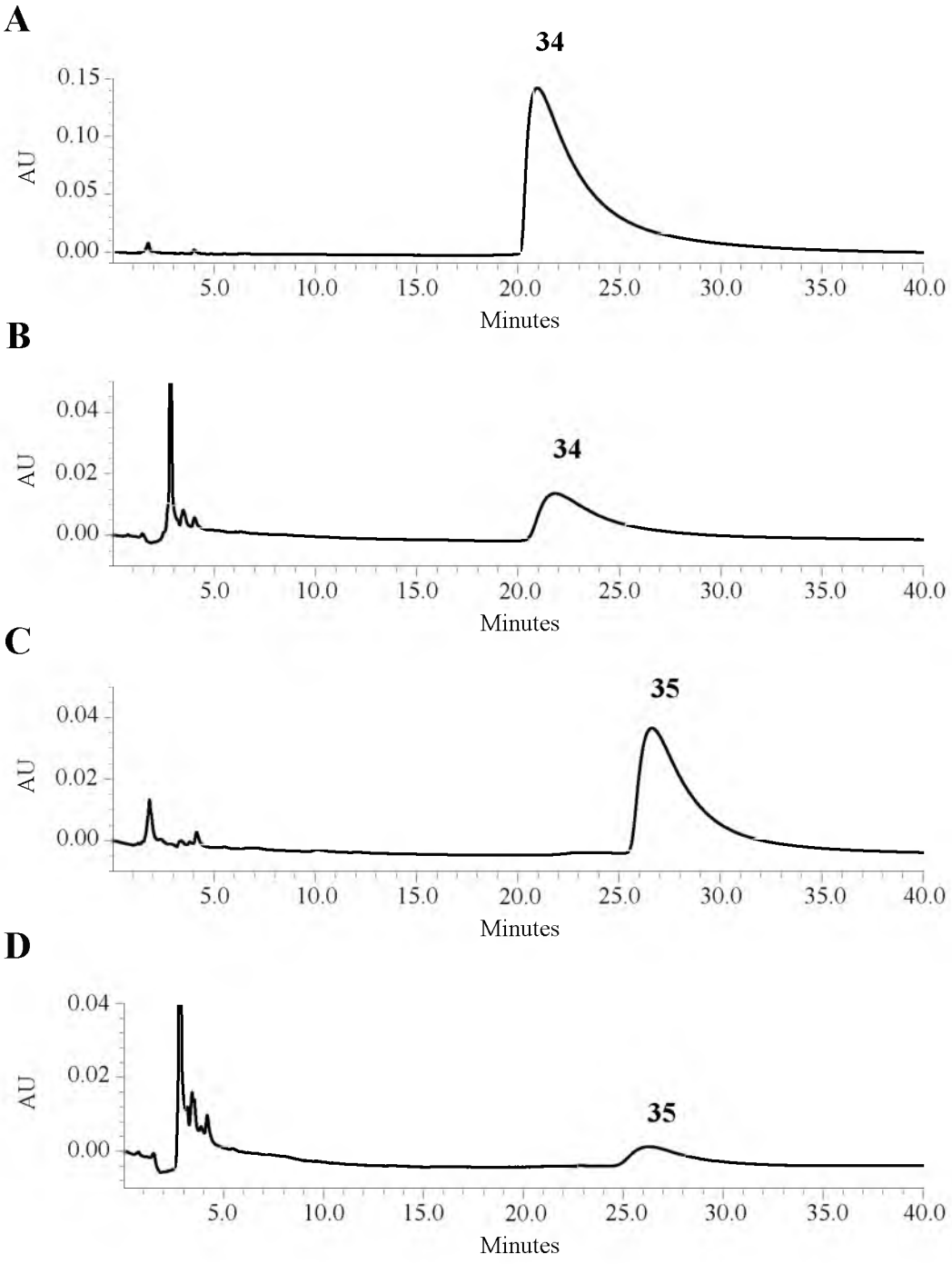


Figure 3.19. HPLC chromatogram of prenylated 4-methyl-L-tryptophan and DMAPP incubation with SirD. A, compound 34 standard; B, compound 34 incubated with SirD; C, compound 35 standard; D, compound 35 incubated with SirD. All traces were monitored at 225 nm.

We suspected that Li and coworkers who previously reported single products for SirD incubations with L-tryptophan and 4-methyl-DL-tryptophan were unable to see the second product peaks due to their use of HPLC solvents containing 0.5% trifluoroacetic acid (TFA).[134,135] HPLC analysis of incubations containing L-tryptophan or **25,** DMAPP, and SirD, in the presence or absence of 0.5% TFA, indicated all products were lost under acidic conditions (data not shown). Therefore, as with kinetic studies, acidic conditions should not be used when analyzing and isolating prenylated aromatic amino acids. A recent report of the disassociation and rearrangement of a dimethylallyl moiety at position N1 of tryptophan-containing cyclic dipeptides in the presence of trichloroacetic acid further supports this premise.[143]

7-Methyl-L-tryptophan (28)

Given the regiospecficity SirD displays by prenylating tryptophans at the N1 and C7 positions of the indole ring, we hypothesized a single prenylation product from enzymatic incubation with 7-methyltryptophan. The substitution of the hydrogen atom at C7 by a methyl group should block alkylation at that position, resulting only in reverse prenylation at N1. As described above, incubation of **28** and DMAPP with SirD gave products **38** (16.3 min, 13%) and **39** (19.6 min, 87%) (Figure 3.20). HRMS analysis indicated monoprenylation of the tryptophan analogue and the UV spectra suggested that the indole nuclei were intact.

The ^{1}H NMR spectrum (Table 3.6) of **38** established that the indole ring contained a dimethylallyl moiety attached in reverse orientation as evidenced by signals for the three vinyl protons at 6.30 ppm (dd, J = 10.8, 17.4 Hz), 4.94 ppm (d, J = 10.8 Hz) and 4.49 ppm (d, J = 17.4 Hz), and the methyl groups at 1.64 ppm (s). The TOCSY spectrum

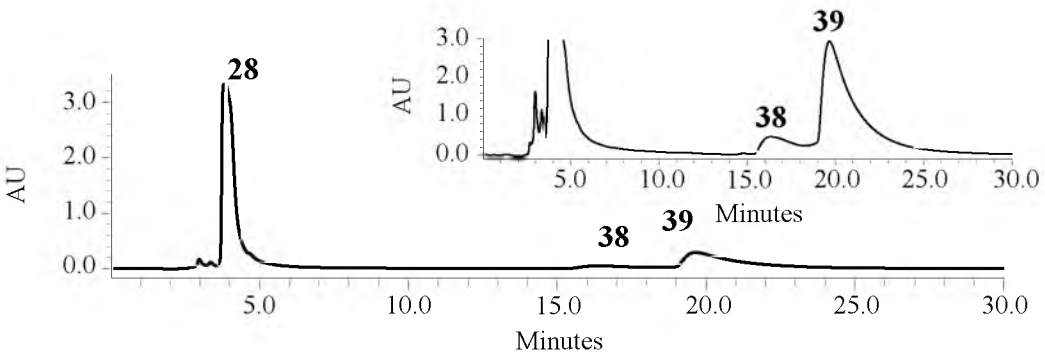
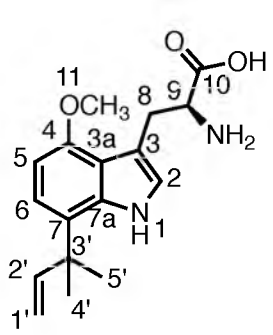
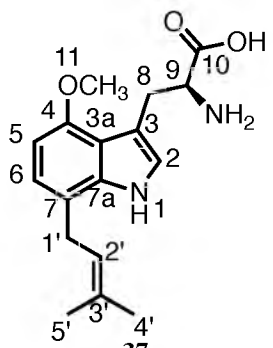
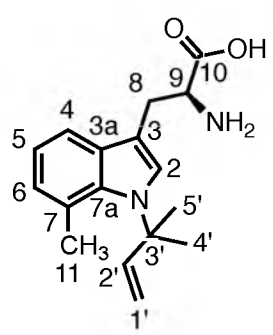
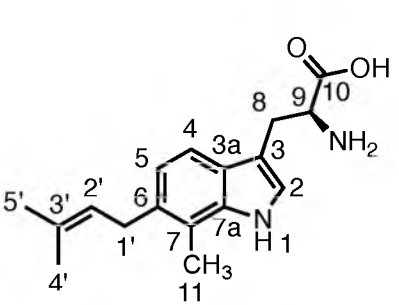


Figure 3.20. HPLC chromatogram of 7-methyl-L-tryptophan and DMAPP incubation with SirD. Product formation was monitored at 225 nm. Inset, identical trace with adjusted y-axis for product peak visualization.

Table 3.6. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from 4-methoxy and 7-methyl-L-tryptophan

Structure/ Compound	 **36**		 **37**		 **38**		 **39**	
Position	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**
1	exc.	-	exc.	-	-	-	exc.	-
2	6.97, s	126.1	7.00, s	126.5	7.43, s	130.3	7.23,s	n.f.
3	-	n.f.	-	n.f.	-	109.5	-	n.f.
3a	-	119.6	-	119.3	-	133.3	-	n.f.
4	-	n.f.	-	154.7	7.46, d, 7.8	119.1	7.48, d, 7.8	118.4
5	6.55, d, 8.4	102.2	6.47, d, 7.8	102.6	7.00, t, 7.8	122.5	7.03, d, 7.8	n.f.
6	7.06, d, 7.8	n.f.	6.84, d, 7.8	n.f.	6.96, d, 6.6	128.5	-	n.f.
7	-	127.8	-	121.5	-	125.5	-	121.6
7a	-	n.f.	-	139.1	-	137.2	-	n.f.
8	3.48, dd, 3.6, 15.0	33.5	3.42, dd, 4.2, 15.0	31.5	3.17-3.23, m	35.1	3.36, dd, 3.6, 15.0	n.f.
	3.02, dd, 8.4, 15.0	-	3.02, dd, 8.4, 15.0	-	2.97-3.04, m	-	3.16, dd, 7.8, 15.0	-
9	3.88, dd, 3.6, 7.8	n.f.	3.85, dd, 4.8, 8.4	59.2	3.66-3.72, m	58.2	3.84, dd, 3.6, 12.0	57.9
10	-	n.f.	-	n.f.	-	n.f.	-	n.f.
11	3.84, s	57.8	3.80, s	57.8	2.54, s	26.6	2.42, s	14.7
1'	5.03, d, 10.8	114.4	3.35-3.37, m	31.2	4.94, d, 10.8	114.7	3.46, d, 7.2	33.8
	4.98, d, 17.4	-	-	-	4.49, d, 17.4	-	-	-
2'	6.00, dd, 10.8, 17.4	149.8	5.30, t, 6.6	124.0	6.30, dd, 10.8, 17.4	150.0	5.32, t, 7.2	125.8
3'	-	42.0	-	137.5	-	62.5	-	135.8
4'	1.35, s	29.4	1.60, s	19.7	1.64, s	34.3	1.77, s	19.7
5'	1.35, s	29.4	1.59, s	27.5	1.64, s	34.3	1.69, s	27.5

δ_H and δ_C in ppm; *J* in Hz; n.f. = not found; exc. = exchanged

showed correlations between the dimethylallyl moiety methyl peaks and the proton at the C2 position of the indole ring (7.43 ppm). The presence of four aromatic protons at C2, C4 (7.46 ppm), C5 (7.00 ppm), and C6 (6.96 ppm) confirms prenylation occurred at the indole nitrogen. Together, these data confirm that SirD catalyzes the reverse prenylation at N1 of 7-methyltryptophan.

The ^{1}H NMR spectrum of **39** had peaks for the protons at C1' (3.46 ppm, d, 7.2 Hz), C2' (5.32 ppm, d, 7.2 Hz), and the methyl groups at C3' (1.77 and 1.69 ppm) for normal prenylation. Signals for the aromatic protons at C2 (7.23, s) and two other coupled aromatic protons ((7.48 ppm, d, 7.8 Hz) and (7.03 ppm, d, 7.8 Hz)) indicated that prenylation occurred at either the 4 or 6 position on the indole ring. The HMBC spectrum showed a correlation between the methyl group at C7 (2.42 ppm) and C3' of the dimethylallyl moiety (135.8 ppm). Although unambiguous structural determination was impossible from the available spectra, a comparison of the ^{1}H NMR spectrum with published data for 4-dimethylallyl-7-methyltryptophan[98] and 7-dimethylallyl-6-methyltryptophan[75] suggests compound **39** is prenylated at the C6 position. 4-Dimethylallyl-7-methyltryptophan reports the protons at C5 (6.99 ppm) and C6 (6.89 ppm). 7-Dimethylallyl-6-methyltryptophan reports the protons at C4 (7.41 ppm) and C5 (6.86 ppm). The aromatic protons of compound **39** correspond well with the latter structure, and more importantly, do not match the chemical shifts of the possible C4-prenylation product. Collectively, these data prove SirD catalyzes the normal prenylation at the C6 position of 7-methyltryptophan.

The consistent reverse prenylation of indole's nitrogen atom for tryptophan, 4-methyltryptophan, and 7-methyltrypophan suggests the binding orientation of these

substrates in relation to the dimethylallyl substrate is similar. The C6 prenylation of 7-methyltryptophan, a substrate blocking one of the preferred prenylation sites, indicates SirD can bind DMAPP in different conformations. A similar prenylation "shift" was reported for IptA, a 6-DMATS.[75] Incubation of 6-methyl-DL-tryptophan with DMAPP and IptA resulted in 7-dimethylallyl-6-methyl-DL-tryptophan. 6-Chloro- and 6-bromo-DL-tryptophan were not substrates.[75] At this point, it is unknown whether SirD would accept 7-halogenated analogues.

4-Methoxy-L-tryptophan (24)

Incubation of **24** with DMAPP and SirD gave products **36** (16.1 min, 8%) and **37** (20.8 min, 92%) (Figure 3.21). HRMS analysis indicated that the molecular ions of the compounds were 68 Da higher than the sodium adduct of **24**, as expected for monoprenylation. The UV spectra for **36** and **37** suggested that the indole nuclei were intact.

The dimethylallyl moiety in compound **36** was attached in a reverse orientation as evidenced by the distinctive signals for the three vinyl protons at 6.00 ppm (dd, J = 10.8, 17.4 Hz), 5.03 ppm (dd, J = 10.8 Hz), and 4.98 ppm (dd, J = 17.4 Hz), and two methyl

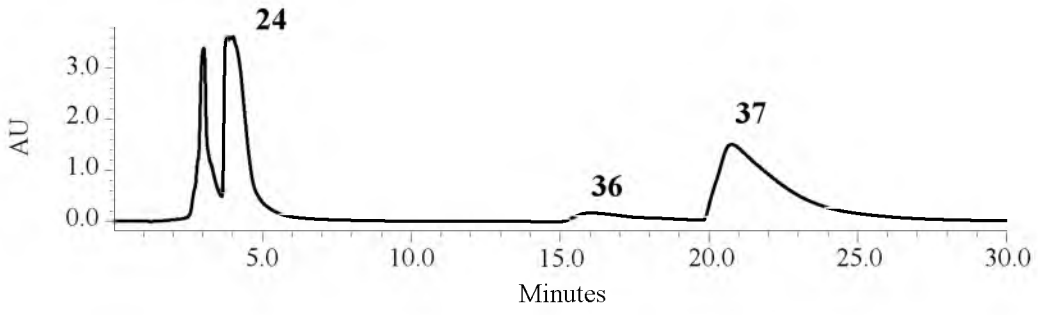


Figure 3.21. HPLC chromatogram of 4-methoxy-L-tryptophan and DMAPP incubation with SirD. Product formation was monitored at 225 nm.

groups at 1.35 ppm. The HMBC spectrum showed correlations between the methyl peaks of the prenyl moiety and C7 (127.8 ppm) in the indole ring. The ROESY spectrum showed correlations between the methyl peaks of the prenyl moiety and the aromatic proton at C6 (7.06 ppm), and between the methoxy substituent at C4 (57.8 ppm) and the proton at C5 (6.55 ppm). Together, these data confirm that **36** is formed by reverse prenylation at the C7 position of 4-methoxytryptophan.

Compound **37** contained a dimethylallyl moiety attached in a normal orientation as evidenced by the signals for the methylene protons at C1' (3.35-3.37 ppm, m), the proton at C2' (5.30 ppm, t, J = 6.6 Hz), and the methyl groups at 1.60 and 1.59 ppm. Signals for the proton at C2 (7.00, s) and two other aromatic protons ((6.84 ppm, d, J = 7.8 Hz) and (6.47 ppm, d, J = 7.8 Hz)) indicated that prenylation had occurred at either the 5 or 7 position of the indole ring. The ROESY spectrum showed a weak correlation between the methylene protons of the prenyl moiety and the aromatic proton at C6 (6.84 ppm). The HMBC spectrum showed correlations between the methylene protons of the prenyl moiety and C7 (121.5 ppm). Collectively, these data confirm that **37** is formed by a normal prenylation at the C7 position of 4-methoxytryptophan.

SirD prenylation of 4-methoxytryptophan was hypothesized to follow the prenylation pattern for tryptophan and 4-methyltryptophan. As expected, SirD prenylated compound **24** at the C7 position of the indole ring in a normal orientation. SirD also catalyzed a reverse prenylation at C7; SirD did not reverse prenylate N1 of 4-methoxytryptophan. Overall, these experiments demonstrate that SirD catalyzes normal or reverse prenylation of tryptophan and tryptophan analogues at three distinct sites on the indole ring, N1, C6, and C7, depending on the substituent (Figure 3.22).

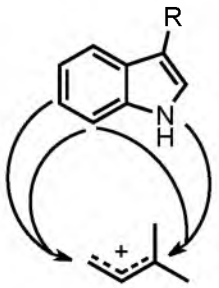


Figure 3.22. Tryptophan prenylation sites and orientations catalyzed by SirD.

The indole ring in tryptophan is susceptible to electrophilic alkylation at N1, at C2, C5, and C7 by direct conjugation with N1 through the benzene ring, and at C3, C4, and C6 by conjugation with N1 through the C2-C3 double bond. The methyl and methoxy substituents introduced at C4 are both electron-donating relative to hydrogen. For tryptophan, alkylation is mostly at C7, with a small amount of N1 alkylation present. For the methyl substituent ($\sigma^+ = 0.31$),[144] alkylation is mostly at N1, along with a smaller amount of C7 alkylation. The more powerful electron-donating methoxy group ($\sigma^+ = 0.78$) directs alkylation exclusively to C7. It is hypothesized that prenylation sites and orientations are determined through a combination of steric and electronic factors.

Investigation into the prenylation of 4-hydroxy-L-tryptophan (**22**) by SirD would help uncover the dominating factor. Being electron-rich yet smaller than **24**, the prenylation product(s) of compound **22** would contribute significantly to this story. Unfortunately, compound **22** is quite sensitive to oxygen; thus a more sophisticated approach must be used to prevent oxidation of it or its prenylated derivatives.

4-Vinyl-L-tryptophan (26) and Dihydroiso-L-tryptophan (29)

Incubation of **26** with DMAPP and SirD yielded products **40** (38.0 min, 52%) and **41** (48.1 min, 48%) (Figure 3.23). UV analysis of the products revealed spectra (λ_{max} =

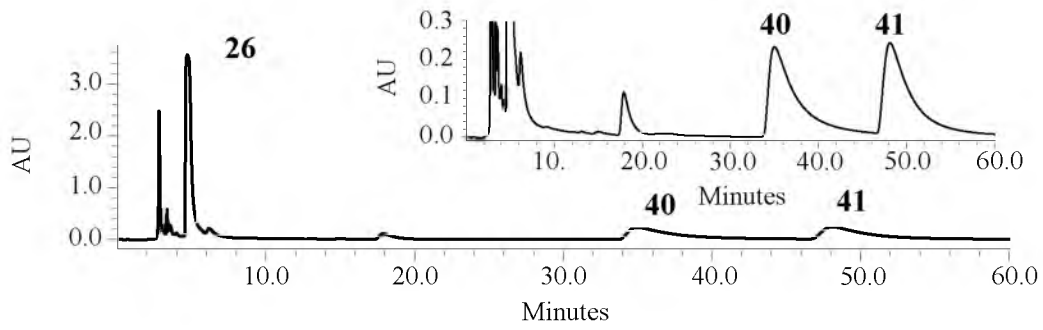


Figure 3.23. HPLC chromatogram of 4-vinyl-L-tryptophan and DMAPP incubation with SirD. Product formation was monitored at 220 nm. Inset, identical trace with adjusted y-axis for product peak visualization.

223.8, 307.0 nm and 219.1, 302.2 nm) very similar to that of the substrate (**26**). Analysis by NMR spectroscopy and MS was not performed, however prenylation is proposed to occur at the N1 and C7 positions of the indole ring. Characterization of the prenylated products would further clarify the prenylation of tryptophan analogues by SirD.

DIT (**29**) provides an interesting substrate analogue as it contains an indoline ring with an amino acid moiety attached on the ring's nitrogen atom. As the amino acid moiety is important for catalysis (Table 3.4), the connectivity of the bicyclic ring to the amino acid moiety may allow new sites to be positioned for prenylation. The 2,3-dihydropyrrole ring with a tertiary amine also prevents prenylation from occurring at the N1, C2 and C3 positions of the indoline ring. Compound **29** was incubated with DMAPP and SirD and one product (**42**, 11.3 min) was formed (Figure 3.24). UV analysis of **42** indicated the indoline ring of DIT (λ_{max} = 205.0, 247.5, 293.8 nm) had changed to what appeared to be an indole ring (λ_{max} = 219.1, 267.7 nm). HRMS analysis revealed that the molecular ion of compound **42** was 68 D higher than DIT indicating the monoprenylation of **29**. The ^{1}H NMR spectrum (see Appendix B NMR 81) also indicated normal prenylation occurred as evidenced by the methylene protons at C1' (3.47 ppm, J = 7.8

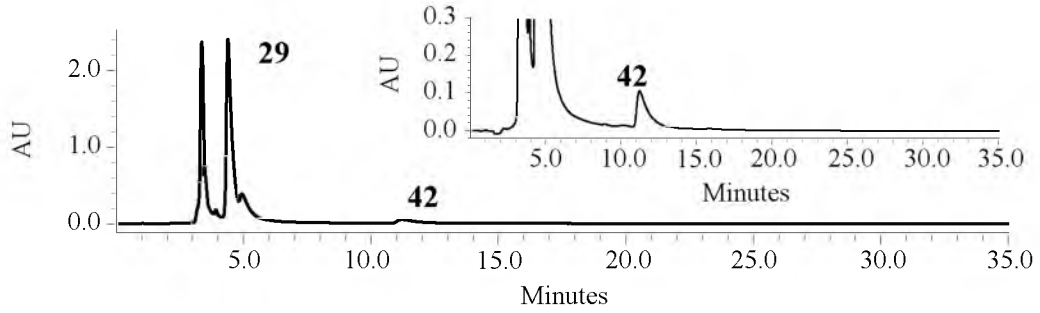


Figure 3.24. HPLC chromatogram of DIT and DMAPP incubation with SirD. Product formation was monitored at 220 nm. Inset, identical trace with adjusted y-axis for product peak visualization.

Hz), the proton at C2' (5.36 ppm, J = 7.8 Hz), and the methyl groups at C3' (1.64 and 1.61 ppm). The presence of four aromatic peaks in a different coupling pattern than DIT indicates that prenylation occurred on the benzene ring. The presence of an aromatic singlet at 7.15 ppm suggests prenylation occurred at C5 or C6. It is unclear, however, what the fourth aromatic peak corresponds to. An intact indole ring, as implied by the UV spectrum, may place the aromatic singlet at a different position. Although at this point, the prenylation site is undetermined, the ability of SirD to prenylate a tryptophan analogue as unique as DIT illustrates its exceptional promiscuity ability.

Overall, the *O*-, *N*-, and *S*-prenylation of tyrosine and its analogues, and the *N*- and *C*-prenylation of tryptophan analogues by SirD, strongly supports an electrophilic alkylation mechanism where substrate orientation within the active site and substituent electronic effects determine the position and type of prenylation. The promiscuity of SirD provides an intriguing argument for the evolution of aromatic prenyltransferases from a common ancestor.

CHAPTER 4

4-DMATS: MULTISITE PRENYLATION OF 4-SUBSTITUTED TRYPTOPHANS

Introduction

Prenylated indole alkaloids constitute a large family of chemically diverse biologically active natural products.[22-24] Biosynthetically, the carbon skeletons of these compounds are constructed from prenyl diphosphates and tryptophan or derivatives of tryptophan. Well-known examples include the ergot alkaloids, the brevianamides, epiamauromine, and verruculogen (Figure 1.1). Collectively a family of enzymes, the dimethylallyl tryptophan synthases (DMATS), catalyze alkylation of the indole ring in tryptophan or tryptophan-containing dipeptides at positions N1, C2, C3, C4, C5, C6, or C7 by dimethylallyl diphosphate (DMAPP) to give the different naturally occurring carbon skeletons.[63,70-76] The dimethylallyl moiety can be attached at C1' ("normal" prenylation) or C3' ("reverse" prenylation), further increasing structural diversity. For wild type DMATSs both normal and reverse prenylation is seen at N1[70,77] and C2.[71,78] Only reverse prenylation is seen at C3 in either the α- or β-configuration by the DMATSs AnaPT[72] or CdpC3PT,[73] respectively. Only normal prenylation is seen at C4,[63,78] C5,[74] and C6.[75] 7-DMATS catalyzes normal prenylation at C7 of L-tryptophan[76] and MpnD catalyzes reverse prenylation at C7 of an indolactam.[79]

4-DMATS catalyzes the normal prenylation of tryptophan at C4 by DMAPP as the first committed step in ergot alkaloid biosynthesis.[30] The enzyme was the first indole prenyltransferase to be purified, characterized, and later available from expression clones and has been studied from *Claviceps purpurea*[62,63,80] and *Aspergillus fumigatus*.[59,78,81,82] 4-DMATS (FgaPT2) from *A. fumigatus* consists of a 10-stranded antiparallel β-barrel surrounded by α-helices.[59] This "ABBA PT" fold is only seen for five other prenyltransferases - FtmPT1[61] and CdpNPT,[60] normal C2 and reverse C3 indole

prenyltransferases, respectively; NphB (formerly Orf2),[56] CloQ,[57] and EpzP,[87] non-indole aromatic prenyltransferases. BLAST alignments with 4-DMATS from *C. purpurea* and other DMATSs report the following sequence identities: 56% with FgaPT2 from *A. fumigatus*, 35% with FtmPT1 from *A. fumigatus*, 35% with CdpC3PT from *Neosartorya fischeri*, 31% with CdpNPT from *A. fumigatus*, 47% with 5-DMATS from *A. clavatus*, 23% with IptA from *Streptomyces* sp. SN-593, and 28% with 7-DMATS from *A. fumigatus*.[69]

The mechanism proposed for prenylation by DMATSs is a dissociative electrophilic alkylation of the indole ring by DMAPP where cleavage of the carbon-oxygen bond in DMAPP gives a dimethylallyl cation-PP_i ion pair, with subsequent alkylation of the indole moiety, followed by loss of a proton to give the prenylated product.[80,131] In the case of 4-DMATS, regiospecific alkylation of C4 of the indole ring by C1' of the allylic cation generates an arenium intermediate, which rearomatizes by deprotonation at C4 to produce dimethylallyl tryptophan (DMAT). The crystal structure of FgaPT2 from *A. fumigatus* in complex with tryptophan and dimethylallyl *S*-thiolodiphosphate, a nonhydrolyzable DMAPP analogue,[88] revealed two active site amino acids thought to be important for catalysis.[59] A hydrogen bond between Glu89 and the indole N-H likely increases the electron density in the indole ring. Lys174 is situated near C4 and is a likely candidate for removing the C4 proton during rearomatization. Recently, a K174A mutant of FgaPT2 gave a C3 reverse-prenylated hexahydropyrroloindole as the major product in addition to a small amount of 4-DMAT.[89] The authors suggested a new mechanism where the initial alkylation is a reverse prenylation at C3 of the indole ring, followed by a Cope rearrangement to give

the normal C4 prenylated arenium intermediate and deprotonation facilitated by Lys174 to produce 4-DMAT.

4-Methyltryptophan has been used as a competitive dead-end analogue of tryptophan in kinetic studies of 4-DMATS.[83] The substitution of the hydrogen atom at the C4 by a methyl group should block alkylation at that position. Previously in the Poulter group, 4-methyl-DL-tryptophan was utilized as a competitive substrate inhibitor with 4-DMATS from *C. purpurea*.[84] During these inhibition studies, a consistently high background of activity was found. It was subsequently discovered that 4-methyltryptophan is an alternate substrate for 4-DMATS and produces four distinct products. In this chapter, we report studies with 4-methyltryptophan and other tryptophan analogues that provide insights about the mechanism of the reaction.

Experimental Procedures

General

NMR experiments were recorded on an INOVA 600 NMR spectrometer equipped with a HCN cryogenic probe and spectra were processed with MestReNova 7.1. All chemical shifts and coupling constants are reported in parts per million (ppm) and hertz (Hz), respectively. ^{1}H chemical shifts are referenced relative to DSS at 0.00 ppm. ^{13}C chemical shifts are referenced relative to DSS at 0.00 ppm. HPLC was conducted on a Water 2690 Separation Module equipped with a Microsorb MV^{TM} C18 5 μm column and a Waters 996 photodiode array detector. Normal phase thin layer chromatography (TLC) was carried out on silica gel 60 glass-backed plates (Merck). Reverse phase TLC was carried out on either C8 glass-backed plates (Silicycle) or C18 glass-backed plates with a

concentrating zone (Merck). Mass spectrometry was performed by Dr. James Muller at the Chemistry Department Mass Spectrometry Laboratory, University of Utah.

Materials

All compounds were obtained from Sigma-Aldrich unless otherwise noted. The following compounds were purchased from the companies indicated: chloramphenicol (MP Biomedicals); DMAPP triammonium salt [1-^{14}C], IPP triammonium salt [1-^{14}C] (American Radiolabeled Chemicals, Inc.); and ampicillin (Roche). Deuterated solvents were purchased from Cambridge Isotope Laboratories, Inc. *E. coli* IDI-1 and pHDMAT were provided by Dr. Jonathan Johnston. 2-Oxo-L-tryptophan and compounds **0393-1** and **0393-2**, the two diastereomers of reverse C3-prenyl-2-oxo-L-tryptophan, were supplied by Dr. Raj Viswanathan.

Bacterial Strains, Plasmids and Culture Conditions

Plasmid pHDMAT is a pET28a construct containing the *C. purpurea* dmaW, which encodes 4-DMATS. Plasmid pGroESL is a construct containing the *E. coli* groE operon[136] and was obtained from Anthony Gatenby. The expression strain *E. coli* BL21 Star™ (DE3) was obtained from Stratagene. Expression strains were grown in liquid LB broth or on solid LB-agar medium at 37 °C. Kanamycin (35 μg/mL) and chloramphenicol (34 μg/mL) were used to select vector-containing *E. coli* strains.

Production and Purification of 4-DMATS

A production strain for soluble 4-DMATS was constructed by cotransforming *E. coli* BL21 Star™ (DE3) with pHDMAT and pGroESL according to the manufacturer's instructions. *E. coli* BL21 Star™ (DE3)/pHDMAT/pGroESL was incubated in 1 L of LB

medium containing 35 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37 °C with shaking at 225 rpm until an OD_{600} of 0.6 was reached. The culture was then cooled to 20 °C and protein expression was induced with IPTG to a final concentration of 0.4 mM. After 18 h, cells were harvested by centrifugation (4,000 x *g*, 20 min, 4 °C). Pelleted cells were resuspended in lysis buffer (50 mM NaH_2PO_4, pH 8.0, containing 300 mM NaCl, 10 mM imidazole, 1 mg/mL of lysozyme and 10 μg/mL of RNase A). After incubation at 0 °C for 30 min, the cell suspension was sonicated for 6 x 10 s at 4 °C. The lysate was clarified by centrifugation (10,000 x *g*, 20 min, 4 °C). 4-DMATS was purified by nickel affinity chromatography (GE Healthcare HisTrap HP) using a 0 – 100% gradient elution of lysis buffer/elution buffer (50 mM NaH_2PO_4, pH 8.0, containing 300 mM NaCl and 500 mM imidazole). The purified protein was dialyzed against 2 x 4 L of dialysis buffer (20 mM Tris-HCl, pH 8.0, containing 1 mM βME), and then against 4 L of dialysis buffer containing 20% glycerol. The protein was flash frozen with liquid nitrogen and stored at -80 °C until use.

Promiscuity Studies

The assays for detecting the promiscuity of aromatic substrates were conducted in 50 mM Tris-HCl, pH 8.0, containing 10 mM $MgCl_2$, 0.1 mg/mL BSA, 10 mM KCl, 1 mM ^{14}C-IPP (0.44-1 mCi/mmol), 1-10 mM aromatic substrate, ~2 μM IDI-1, and 0.65-9.2 μM 4-DMATS in a total volume of 50 μL. For thioanalogue assays, 20 mM βME was included to reduce unwanted disulfides. Typically, each reaction mixture was incubated at 30 °C for 1 h and 20 h. After incubation, a 10 μL portion of assay mixture was spotted and developed by TLC. For normal phase TLC, a solvent system of 1:9 acetic acid/EtOH or 2:8 H_2O/ACN was used to develop the plates. Reverse phase C18-

TLC was performed as described in Chapter 3. The developed TLC plates were then imaged on a storage phosphor screen (Molecular Dynamics), scanned by a Typhoon 8600 Variable Mode Imager (GE Healthcare), and the data were visualized and processed with ImageQuant 5.2.

Kinetic Studies

All kinetic assays were conducted in 50 mM Tris-HCl buffer, pH 8.0, containing 4 mM $MgCl_2$ and 50 μM ^{14}C-DMAPP (10 μCi/μmol), in a total volume of 100 μL. Each reaction was incubated at 30 °C and quenched by heating at 100 °C for 30 s. After centrifugation, 10 μL of reaction mixture were spotted and developed by RP-C8 TLC using a solvent system of 4:6 25 mM NH_4HCO_3/methanol. TLC data visualization/processing was peformed as previously described. Background radioactivity was determined from incubations without the tryptophan substrate. Each kinetic assay was performed in duplicate or triplicate. For L-tryptophan, the assay contained 50 nM 4-DMATS and L-tryptophan concentrations of 3, 10, 20, 30 and 60 μM, with a 5 min incubation. For the tryptophan analgoues, the assays contained 500 nM 4-DMATS with 20 min incubations. For 4-methyltryptophan, final concentrations were 30, 100, 300 and 600 μM; for 4-methoxytryptophan, 30, 100, 300, 600, 1000 and 3000 μM; and for 4-aminotryptophan, 3, 10, 20, 30, 60, and 100 μM.

Product Studies

Incubations were in 50 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$, 8.8 mM DMAPP, 20 mM tryptophan analogue and 15 μM 4-DMATS in a total volume of 5 mL. Glycerol present in the storage buffer was removed by repeated

dialysis/centrifugation using an Amicon Ultra 10 kDa MWCO filter. The assay mixture was incubated at 30 °C for 18-42 h. 4-DMATS was removed by 10 kDa MWCO filtration, samples were concentrated by lyophilization, and the products were purified by C18 RP-HPLC at a flow rate of 1 mL/min with isocratic elution using ACN:H_2O (25/75). Products were detected with simultaneous monitoring at 206, 225 and 254 nm. Product ratios were calculated by integrating the area under each product peak visualized at 214 nm (4-methyl/4-methoxy) or 225 nm (4-amino/DIT).

Structure of 4-Substituted-dimethylallyltryptophan Products

After purification of the prenylated product(s), each product was lyophilized, dissolved in D_2O (D, 99.9%), lyophilized again, dissolved in D_2O (D, 99.96%) and placed in a Norell 3 mm NMR tube (Sigma). Unless otherwise noted, the following NMR experiments were performed: 1D ^1H, COSY ^1H-^1H, TOCSY ^1H-^1H, ROESY ^1H-^1H, HSQC or HMQC ^1H-^{13}C, and HMBC ^1H-^{13}C. All spectra were processed with MestReNova 7.1.

(2*S*)-4-Methyl-3a-(3-methylbut-1-en-3-yl)-1,2,3,3a,8,8a-hexa-hydropyrrolo[2,3-*b*]indole-2-carboxylic Acid (43)

^1H NMR (600 MHz, D_2O) δ 0.85 (s, 3H), 0.93 (s, 3H), 2.29 (s, 3H), 2.34 (dd, J = 8.4, 13.8 Hz, 1H), 2.72 (dd, J = 7.8, 13.8 Hz, 1H), 3.49 (t, J = 7.8 Hz, 1H), 4.94 (dd, J = 0.6, 17.4 Hz, 1H), 4.95 (dd, J = 0.6, 10.2 Hz, 1H), 5.06 (s, 1H), 5.85 (dd, J = 10.8, 17.4 Hz, 1H), 6.48 (d, J = 7.8 Hz, 1H), 6.61 (d, J = 7.8 Hz, 1H), 7.00 (t, J = 7.8 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 23.1, 25.7, 26.1, 38.0, 44.7, 63.5, 68.1, 83.3, 110.8, 116.2, 125.8, 130.9, 131.5, 138.9, 147.6, 152.7, 179.5; UV (ACN:H_2O 1:3) λ_{max} 209.7, 242.8,

293.8 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{17}H_{22}N_2O_2Na$ 309.1579, found 309.1581.

(2*S*)-4-Methyl-3a-(2-methylbut-2-en-4-yl)-1,2,3,3a,8,8a-hexahydropyrrolo[2,3-*b*]indole-2-carboxylic Acid (44)

^1H NMR (600 MHz, D_2O) δ 1.46 (s, 3H), 1.47 (s, 3H), 2.00 (dd, *J* = 8.4, 12.6 Hz, 1H), 2.22 (s, 3H), 2.34 (dd, *J* = 8.4, 15.0 Hz, 1H), 2.51 (dd, *J* = 6.6, 12.6 Hz, 1H), 2.55 (dd, *J* = 6.6, 15.0 Hz, 1H), 3.40 (m, 1H), 4.77 (t, *J* = 7.2, 1H), 4.84 (s, 1H), 6.45 (d, *J* = 7.8 Hz, 1H), 6.57 (d, *J* = 7.8 Hz, 1H), 6.94 (t, *J* = 7.8 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.8, 20.0, 27.4, 37.4, 44.2, 61.1, 63.4, 84.2, 110.5, 121.7, 124.5, 130.8, 133.7, 137.7, 138.6, 152.2; UV (ACN:H_2O 1:3) λ_{max} 206.2, 239.2, 289.1, 293.8 nm; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{17}H_{23}N_2O_2$ 287.1760, found 287.1755.

***N*1-(2-Methylbut-2-en-4-yl)-4-methyl-L-tryptophan (46, Normal *N*-Dimethylallyl-4-methyl-L-tryptophan)**

^1H NMR (600 MHz, D_2O) δ 1.60 (s, 3H), 1.70 (s, 3H), 2.58 (s, 3H), 2.98 (dd, *J* = 9.0, 15.0 Hz, 1H), 3.42 (dd, *J* = 4.8, 15.0 Hz, 1H), 3.61 (dd, *J* = 4.8, 9.0 Hz, 1H), 4.60 (d, *J* = 6.6, 2H), 5.24 (t, *J* = 6.6 Hz, 1H), 6.80 (d, *J* = 7.2 Hz, 1H), 7.03 (dd, *J* = 7.2, 8.4 Hz, 1H), 7.07, (s, 1H), 7.19 (d, *J* = 8.4 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.8, 22.2, 27.2, 32.4, 46.3, 59.6, 110.7, 112.0, 121.9, 123.3, 124.4, 128.2, 130.6, 133.8, 138.9, 140.4; UV (ACN:H_2O 1:3) λ_{max} 225.0, 276.0, 290.3 nm; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{17}H_{23}N_2O_2$ 287.1760, found 287.1760.

5-(2-Methylbut-2-en-4-yl)-4-methoxy-L-tryptophan (49, Normal 5-Dimethylallyl-4-methoxy-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.60 (s, 3H), 1.67 (s, 3H), 3.06 (dd, *J* = 9.0, 15.0 Hz, 1H), 3.35 (d, *J* = 7.2 Hz, 2H), 3.41-3.44 (m, 1H), 3.79 (s, 3H), 3.87 (dd, *J* = 4.8, 9.0 Hz, 1H), 5.25 (t, *J* = 7.2 Hz, 1H), 6.98 (d, *J* = 8.4 Hz, 1H), 7.10 (s, 1H), 7.15 (d, *J* = 8.4 Hz, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.8, 27.5, 29.7, 65.1, 111.7, 126.2, 126.8, 127.8, 136.3; UV (ACN:H_2O 1:3) λ_{max} 222.7, 270.0 nm; HRMS-ESI TOF *m/z* $[M + Na]^+$ calcd for $C_{17}H_{22}N_2O_3Na$ 325.1528, found 325.1530.

5-(2-Methylbut-2-en-4-yl)-4-amino-L-tryptophan (51, Normal 5-Dimethylallyl-4-amino-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.61 (s, 3H), 1.66 (s, 3H), 3.17 (dd, *J* = 8.4, 15.6 Hz, 1H), 3.24 (d, *J* = 7.2 Hz, 2H), 3.42-3.47 (m, 1H), 3.77-3.82 (m, 1H), 5.19 (t, *J* = 7.2 Hz, 1H), 6.87 (d, *J* = 8.4 Hz, 1H), 6.89 (d, *J* = 8.4 Hz, 1H), 7.02 (s, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.6, 27.4, 31.0, 31.9, 58.7, 107.0, 109.9, 120.0, 125.0, 127.0, 127.3, 136.8, 138.4, 139.4; UV (ACN:H_2O 1:3) λ_{max} 225.0, 271.2, 292.7 nm; HRMS-ESI TOF *m/z* $[M + H]^+$ calcd for $C_{16}H_{22}N_3O_2$ 288.1712, found 288.1715.

7-(2-Methylbut-2-en-4-yl)-4-amino-L-tryptophan (52, Normal 7-Dimethylallyl-4-amino-L-tryptophan)

^{1}H NMR (600 MHz, D_2O) δ 1.61 (s, 3H), 1.62 (s, 3H), 3.15 (dd, *J* = 7.2, 15.0 Hz, 1H), 3.35 (d, *J* = 7.8, 2H), 3.34-3.40 (m, 1H), 3.68-3.72 (m, 1H), 5.31 (t, *J* = 7.8 Hz, 1H), 6.36 (d, *J* = 7.2 Hz, 1H), 6.75 (d, *J* = 7.2 Hz, 1H), 7.05 (s, 1H); ^{13}C NMR (150 MHz, D_2O) δ 19.4, 27.3, 31.2, 31.4, 58.9, 110.1, 111.3, 119.9, 120.3, 124.1, 124.2, 126.6,

137.3, 138.9, 139.9; UV (ACN:H_2O 1:3) λ_{max} 223.8, 273.6 nm; HRMS-ESI TOF *m/z* [M + H]$^+$ calcd for $C_{16}H_{22}N_3O_2$ 288.1712, found 288.1717.

Rearrangement Studies

The assay for the detection of prenyl rearrangement was conducted in 50 mM Tris-HCl, pH 8.0, containing 210 nM 4-DMATS and 5 mM of compound **0393-1** or **0393-2** in a total volume of 0.5 mL. Additional assays included 2 mM DMAPP or 2 mM PP_i. The control reaction was incubated in the presence of 2 mM DMAPP and 2 mM PP_i and in absence of 4-DMATS. The assay mixture was incubated at 30 °C for 18 h. 4-DMATS was removed by 10 kDa MWCO filtration and the filtrates were analyzed by HPLC as described above. The detection of aromatic products was carried out at 254 nm.

Results and Discussion

pHDMAT/pGroESL Cloning, Expression and Purification

As established by Wang and Poulter, soluble 4-DMATS from *C. purpurea* was obtained through co-expression with the chaperone protein complex GroEL/ES.[84] Co-transformation of the pHDMAT (His-tagged 4-DMATS) and pGroELS (GroE operon) plasmids and subsequent heterologous overexpression gave soluble 4-DMATS. Purification was achieved through nickel-affinity chromatography and SDS-PAGE gave a band at ~55 kDa (Figure 4.1), near the calculated weight of 54,021 Da.

Promiscuity Studies

Commercially available and synthetic substrate analogues of aromatic amino acids were tested as alternate substrates in the presence of DMAPP and 4-DMATS by the qualitative radio-TLC coupled assay outlined in Chapter 3 (Figure 4.2). 4-DMATS

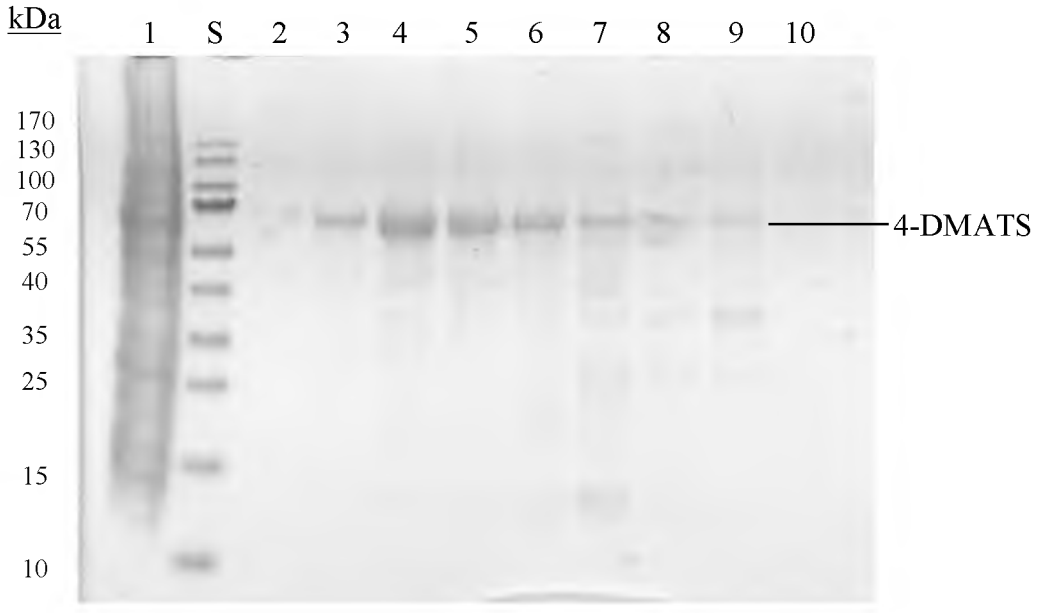


Figure 4.1. SDS-PAGE of pHDMAT/pGroESL co-expression in BL21 (DE3). Lane 1, HisTrap column flow-through; lane S, molecular weight standards; lanes 2-10, HisTrap column elutions.

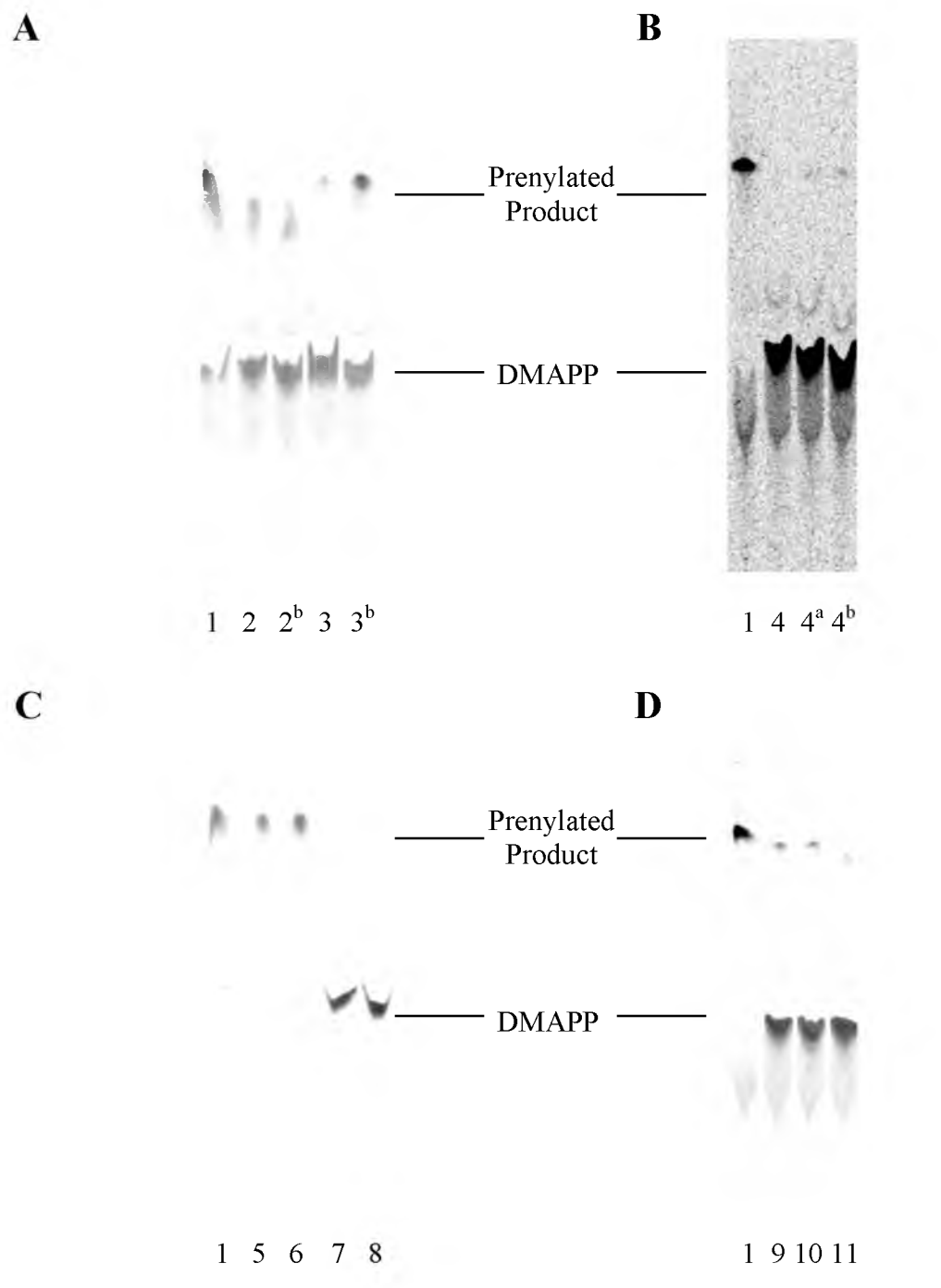


Figure 4.2. Radioautography TLCs of substrate promiscuity assays for 4-DMATS. Lanes 1, L-tryptophan positive control; lanes 2, 4-amino-L-tryptophan; lanes 3, dihydroiso-L-tryptophan; lanes 4, 4-vinyl-L-tryptophan; lane 5, 4-hydroxy-L-trytophan; lane 6, 7-hydroxy-L-tryptophan; lane 7, 4-hydroxyindole; lane 8, 7-hydroxyindole; lane 9, 2-methyl-L-tryptophan; lane 10, 4-mercapto-L-tryptophan; lane 11, 2-oxo-L-tryptophan. Superscripts designate aromatic substrate concentrations: a, 5 mM; b, 10 mM; no superscript, 1 mM.

accepted tryptophan analogues, including analogues containing substituents at the 4 position of the indole ring. The following tryptophan analogues were found to be substrates of 4-DMATS: 4-amino (**21**), 4-hydroxy (**22**), 4-mercapto (**23**), 4-methoxy (**24**), 4-methyl (**25**), 4-vinyl (**26**), 7-hydroxy (**27**), 2-methyl (**20**), 2-oxotryptophan, and dihydroisotryptophan (DIT, **29**). Some substrates showed turnover at concentrations as low as 1 mM (**21** and **29**) while others only showed turnover at concentrations of 10 mM (**24** and **25**) (Figure 4.2). The products of the analogues not substituted at position C4 of the indole ring were hypothesized to be normal C4 prenylated. The 4-substituted analogues were hypothesized to be normal prenylated at the heteroatom of compounds **21**, **22**, and **23** or another available position on the indole ring for compounds **24, 25** and **26**. 4-DMATS may prenylate 4-substituted analogues at the C5 position in a normal orientation. This prenylation shift was shown by SirD prenylating 7-methyltryptophan at the C6 position (Chapter 3). Other aromatic analogues determined not to be substrates for 4-DMATS include L-tyrosine, 4-vinyl-L-phenylalanine (**15**), 4- and 7-hydroxyindoles and indene (Table 4.1, TLC data not shown). Similar to SirD, the inability of 4-DMATS to prenylate indole analogues such as 4- and 7-hydroxyindole indicates the importance for the aromatic substrate to contain an amino acid moiety.

Kinetic Studies

Michaelis-Menten kinetic parameters for tryptophan and its 4-substituted analogues were determined (Figure 4.3, Table 4.2). Rates (k_{cat}) and Michaelis constants (K_M) were determined from a nonlinear regression fit by GraFit 5.0.11. K_M = 20 μM for L-tryptophan, which compared well with the previously reported value.[80] K_M = 17 μM for 4-aminotryptophan was comparable to the normal substrate, while those for the 4-methyl-

Table 4.1. Aromatic substrate analogues tested for turnover with 4-DMATS and the corresponding product ratios

Compound	[Compound] mM	Reaction (y/n)	Product Ratios (%)
L-Tryptophan	1	y	100
4-Amino-L-tryptophan[a] (**21**)	1, 10	y, y	82:18
4-Hydroxy-L-tryptophan (**22**)	2.5, 5, 10	y, y, y	100
4-Mercapto-L-tryptophan (**23**)	1	y	n.d.
4-Methoxy-L-tryptophan[b] (**24**)	2.5, 5, 10	n, n, y	5:7:71:17
4-Methyl-L-tryptophan[b] (**25**)	1, 10	n, y	44:5:7:44
4-Vinyl-L-tryptophan (**26**)	1, 5, 10	n, y, y	n.d.
7-Hydroxy-L-tryptophan (**27**)	2.5, 5, 10	y, y, y	n.d.
2-Methyl-L-tryptophan (**20**)	1	y	n.d.
Dihydroiso-L-tryptophan (**29**)	1, 10	y, y	36:64
2-Oxo-L-tryptophan	1	y	100
4-Hydroxyindole	4	n	-
7-Hydroxyindole	4	n	-
Indene	10	n	-
L-Tyrosine	1	n	-
4-Vinyl-L-phenylalanine (**15**)	5, 10	n, n	-

[a]Product ratios determined by peak integration at 225 nm
[b]Product ratios determined by peak integration at 214 nm
n.d., not determined

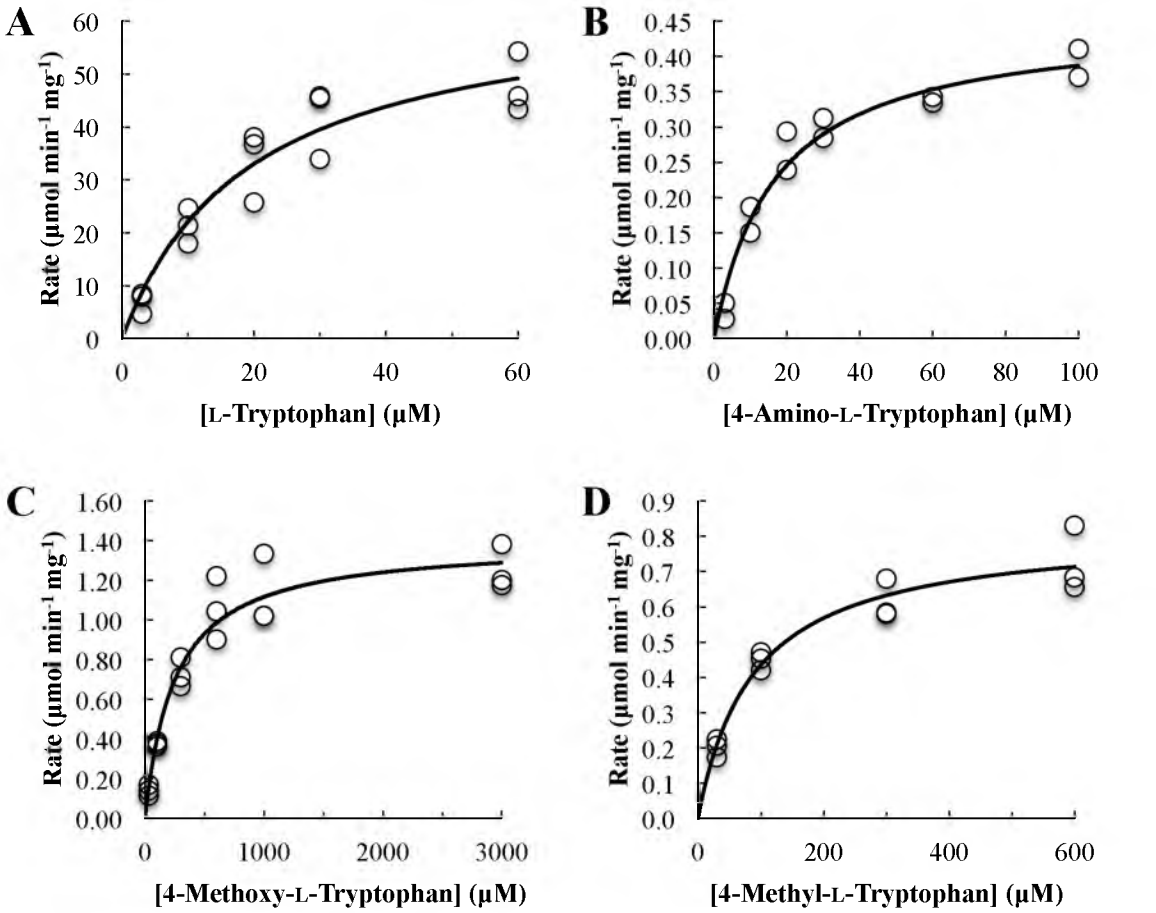


Figure 4.3. Michaelis-Menten kinetic plots for select 4-DMATS substrates. Each graph plots substrate concentration vs. rate and illustrates individual data points and a nonlinear regression curve. A, L-tryptophan; B, 4-amino-L-tryptophan; C, 4-methoxy-L-tryptophan; D, 4-methyl-L-tryptophan. All kinetic experiments were conducted at 30 °C.

Table 4.2. Kinetic parameters for 4-DMATS substrates at 30 °C

Substrate	V_{max} (µmol min^{-1} mg^{-1})	K_M (µM)	k_{cat} (s^{-1})	k_{cat} / K_M (s^{-1} M^{-1})
L-TRP	65 ± 7	20 ± 5	60 ± 6	3.1 x 10^6
4-NH_2-L-TRP (**21**)	0.45 ± 0.03	17 ± 3	0.42 ± 0.02	2.5 x 10^4
4-OCH_3-L-TRP (**24**)	1.4 ± 0.1	250 ± 42	1.3 ± 0.1	5.1 x 10^3
4-CH_3-L-TRP (**25**)	0.82 ± 0.04	89 ± 16	0.75 ± 0.04	8.5 x 10^3

and 4-methoxy analogues, 89 and 250 μM, respectively, were substantially higher. k_{cat} for the tryptophan analogues ranged between 0.42 and 1.3 s^{-1}, substantially smaller than k_{cat} = 60 s^{-1} for tryptophan. The catalytic efficiencies (k_{cat} / K_M) for the analogues were less than 1% of the normal reaction.

Product Studies

Previously in the Poulter group, racemic 4-methyltryptophan was used as a substrate with *C. purpurea* 4-DMATS.[84] However, in a study of the *A. fumigatus* enzyme (FgaPT2) D-tryptophan was also a substrate, although at a rate of only 1.8% that of the L-enantiomer.[78] Thus, we decided to use enantiomerically pure 4-methyl-L-tryptophan (**25**) in our study. The substrate was synthesized as described in Chapter 2. The same procedure was used to synthesize all tryptophan analogues used in this chapter.

4-Methyltryptophan (25)

Incubation of **25** and DMAPP with 4-DMATS gave four products (Figure 4.4). The UV spectra of compounds **43** (11.4 min, 44%) and **44** (15.1 min, 5%) had peaks at 209.7, 242.8, 293.8 nm and 206.2, 239.2, 289.1, 293.8 nm, respectively, suggesting that the indole chromophore ($\lambda_{max} \approx$ 278 nm) had been disrupted to give an indoline-like structure.[145] Compounds **45** (24.6 min, 7%) and **46** (26.9 min, 44%) had peaks at 223.8, 272.4 nm and 225.0, 276.0, 290.3 nm, respectively, suggesting that they contained intact indole chromophores. Chromatography of the reaction mixture from **25** on a C18 RP-HPLC column with an acidic solvent system (1% TFA) resulted in complete loss of **43** while the ratios for compounds **44**, **45** and **46** did not change (data not shown). HRMS analysis of the four products gave molecular ions that were 68 Da higher than the

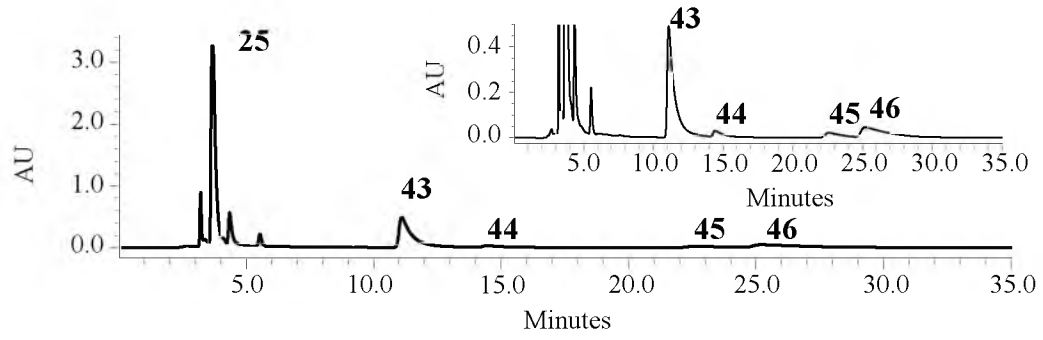
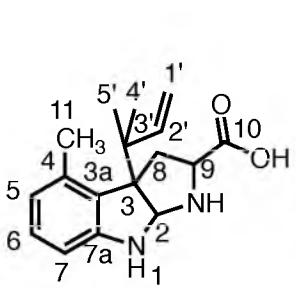
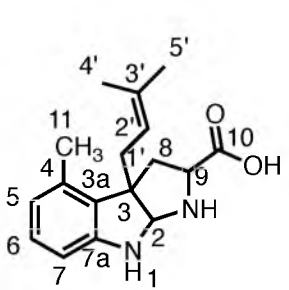
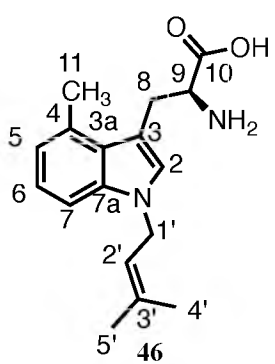
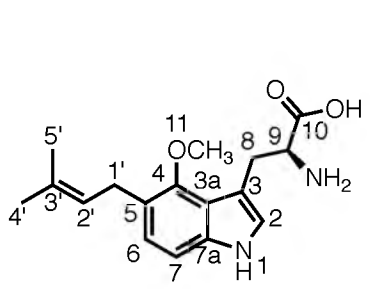


Figure 4.4. HPLC chromatogram of 4-methyl-L-tryptophan and DMAPP incubation with 4-DMATS. Product formation was monitored at 254 nm. Inset, identical trace with adjusted y-axis for product peak visualization.

corresponding peaks for **25** or its sodium adduct, indicating the addition of one isoprene unit.

The structures of the products were determined from their ^{1}H and ^{13}C NMR spectra (Table 4.3). The dimethylallyl moiety in compound **43** was attached in a reverse orientation as evidenced by the distinctive signals for the three vinyl protons at 5.85 ppm (dd, J = 10.8, 17.4 Hz), 4.95 ppm (dd, J = 0.6, 10.2 Hz), and 4.94 ppm (dd, J = 0.6, 17.4 Hz), and two methyl groups at 0.93 and 0.85 ppm. The HMBC spectrum showed correlations between the methyl peaks of the prenyl moiety and C3' (44.7 ppm) of the isoprene unit and C3 (68.1 ppm) in the indoline moiety. Aromatic protons at C5 (6.61 ppm), C6 (7.00 ppm), and C7 (6.48 ppm) support prenylation at C3 of the indole ring. In addition, the peak for the proton at C2 (5.06 ppm) is near reported values for C2 diamine protons in tricyclic C3-prenylated tryptophan derivatives.[60,89] Collectively, these data confirm that **43** is formed by a reverse prenylation at the C3 position of 4-methyltryptophan, followed by cyclization to give a hexahydropyrroloindole (HHPI) structure.

Table 4.3. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from 4-methyl and 4-methoxy-L-tryptophan

Structure/ Compound	 **43**		 **44**		 **46**		 **49**	
Position	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**
1	exc.	-	exc.	-	-	-	exc.	-
2	5.06, s	83.3	4.84, s	84.2	7.07, s	130.6	7.10, s	127.8
3	-	68.1	-	61.1	-	112.0	-	n.f.
3a	-	138.9	-	137.7	-	128.2	-	n.f.
4	-	130.9	-	133.7	-	133.8	-	n.f.
5	6.61, d, 7.8	125.8	6.57, d, 7.8	124.5	6.80, d, 7.2	123.3	-	n.f.
6	7.00, t, 7.8	131.5	6.94, t, 7.8	130.8	7.03, dd, 7.2, 8.4	124.4	6.98, d, 8.4	126.8
7	6.48, d, 7.8	110.8	6.45, d, 7.8	110.5	7.19, d, 8.4	110.7	7.15, d, 8.4	111.7
7a	-	152.7	-	152.2	-	138.9	-	n.f.
8	2.72, dd, 7.8, 13.8	38.0	2.51, dd, 6.6, 12.6	44.2	3.42, dd, 4.8, 15.0	32.4	3.41-3.44, m	n.f.
	2.34, dd, 8.4, 13.8	-	2.00, dd, 8.4, 12.6	-	2.98, dd, 9.0, 15.0	-	3.06, dd, 9.0, 15.0	-
9	3.49, t, 7.8	63.5	3.40, m	63.4	3.61, dd, 4.8, 9.0	59.6	3.87, dd, 4.8, 9.0	n.f.
10	-	179.5	-	n.f.	-	n.f.	-	n.f.
11	2.29, s	23.1	2.22, s	20.0	2.58, s	22.2	3.79, s	65.1
1'	4.95, dd, 0.6, 10.2	116.2	2.55, dd, 6.6, 15.0	37.4	4.60, d, 6.6	46.3	3.35, d, 7.2	29.7
	4.94, dd, 0.6, 17.4	-	2.34, dd, 8.4, 15.0	-	-	-	-	-
2'	5.85, dd, 10.8, 17.4	147.6	4.77, t, 7.2	121.7	5.24, t, 6.6	121.9	5.25, t, 7.2	126.2
3'	-	44.7	-	138.6	-	140.4	-	136.3
4'	0.93, s	26.1	1.47, s	19.8	1.70, s	19.8	1.67, s	19.8
5'	0.85, s	25.7	1.46, s	27.4	1.60, s	27.2	1.60, s	27.5

δ_H and δ_C in ppm; *J* in Hz; n.f. = not found; exc. = exchanged

Compound **44** also contains an HHPI moiety formed by normal prenylation at C3 of 4-methyltryptophan. Normal prenylation was confirmed by the presence of proton signals for methylene protons at C1' (2.34 (dd, J = 8.4, 15.0 Hz) and 2.55 ppm (dd, J = 6.6, 15.0 Hz)), the proton at C2' (4.77 ppm, app t, $J \approx 7.2$ Hz), and the methyl groups at 1.47 and 1.46 ppm. Cross peaks were also seen between a proton at C1' (2.34 ppm) and C3 (61.1 ppm) and between the proton at C2 (4.84 ppm) and C1' (37.4 ppm) in the HMBC spectrum. In addition, signals were seen for the aromatic protons at C5 (6.57 ppm), C6 (6.94 ppm), and C7 (6.45 ppm). Thus, 4-DMATS catalyzes the normal and reverse prenylation at the C3 of 4-methyltryptophan. Very recently, FtmPT1 from *A. fumigatus* was seen to catalyze a normal prenylation at the C3 position of cyclic dipeptides.[146]

The ^{1}H spectrum of **46** had peaks for the protons at C1' (4.60 ppm, d, J = 6.6 Hz) and C2' (5.24 ppm, t, J = 6.6 Hz) and the methyl groups at C3' (1.70 and 1.60 ppm) for normal prenylation. Signals for aromatic protons at C5 (6.80 ppm), C6 (7.03 ppm), and C7 (7.19 ppm) indicated that prenylation had occurred in the pyrrole ring. The chemical shifts of C1' (46.3 ppm) and the directly attached protons (4.60 ppm), along with cross peaks between the C1' protons and C2 (130.6 ppm) in the HMBC spectrum indicate that normal prenylation occurred at the N1 position on the indole ring. The UV spectrum and the resonance at 7.07 ppm for the proton at C2 indicated that the indole ring was intact. This agreed with previous experimental data from our laboratory. A negative ion fast atom bombardment mass spectroscopy spectrum of fully exchanged compound **46** in deuterated glycerol and D_2O verified indole ring nitrogen substitution.[84] Unfortunately, we were unable to obtain sufficient quantities of **45** to establish its structure, although as

mentioned earlier, the indole ring appeared to be intact. Co-injection experiments with **45** and two products isolated and characterized in Chapter 3, normal 7-dimethylallyl-4-methyl-L-tryptophan (**35**) and reverse *N*-dimethylallyl-4-methyl-L-tryptophan (**34**), excluded these structures as candidates for **45** (Figure 4.5).

4-Methoxytryptophan (24)

As described for 4-methyltryptophan, incubation of **24** with 4-DMATS gave products **47** (11.8 min, 5%), **48** (13.3 min, 7%), **49** (17.5 min, 71%), and **50** (21.4 min, 17%) (Figure 4.6). HRMS analysis indicated that the molecular ions of the compounds were 68 Da higher than **24** and its sodium adduct as expected for monoprenylation. The UV spectra for **47** and **48** suggested that the indole moiety had been disrupted, while those for **49** and **50** indicated that the indole nucleus was intact.

The ^{1}H NMR spectrum (Table 4.3) of **49** had peaks for the protons at C1' (3.35 ppm, d, J = 7.2 Hz) and C2' (5.25 ppm, t, J = 7.2 Hz) and the methyl groups at C3' (1.67 and 1.60 ppm) for normal prenylation. Signals for the proton at C2 (7.10, s) and two other aromatic protons ((7.15 ppm, d, J = 8.4 Hz) and (6.98 ppm, d, J = 8.4 Hz)) indicated that prenylation had occurred at either the 5 or 7 position of the indole ring. Although we were unable to unambiguously determine the prenylation site from these data, we previously isolated and characterized normal 7-dimethylallyl-4-methoxy-L-tryptophan (**37**, Chapter 3). A comparison of UV spectra and NMR chemical shifts for both compounds confirm they are distinct products implying **49** is prenylated at C5. Thus, 4-DMATS catalyzes a normal prenylation at C5 of 4-methoxytryptophan.

Given the elution and UV profiles for compounds **47** – **50** were similar to those of the prenylated 4-methyltryptophan derivatives, it is likely **47** and **48** are indolines normal

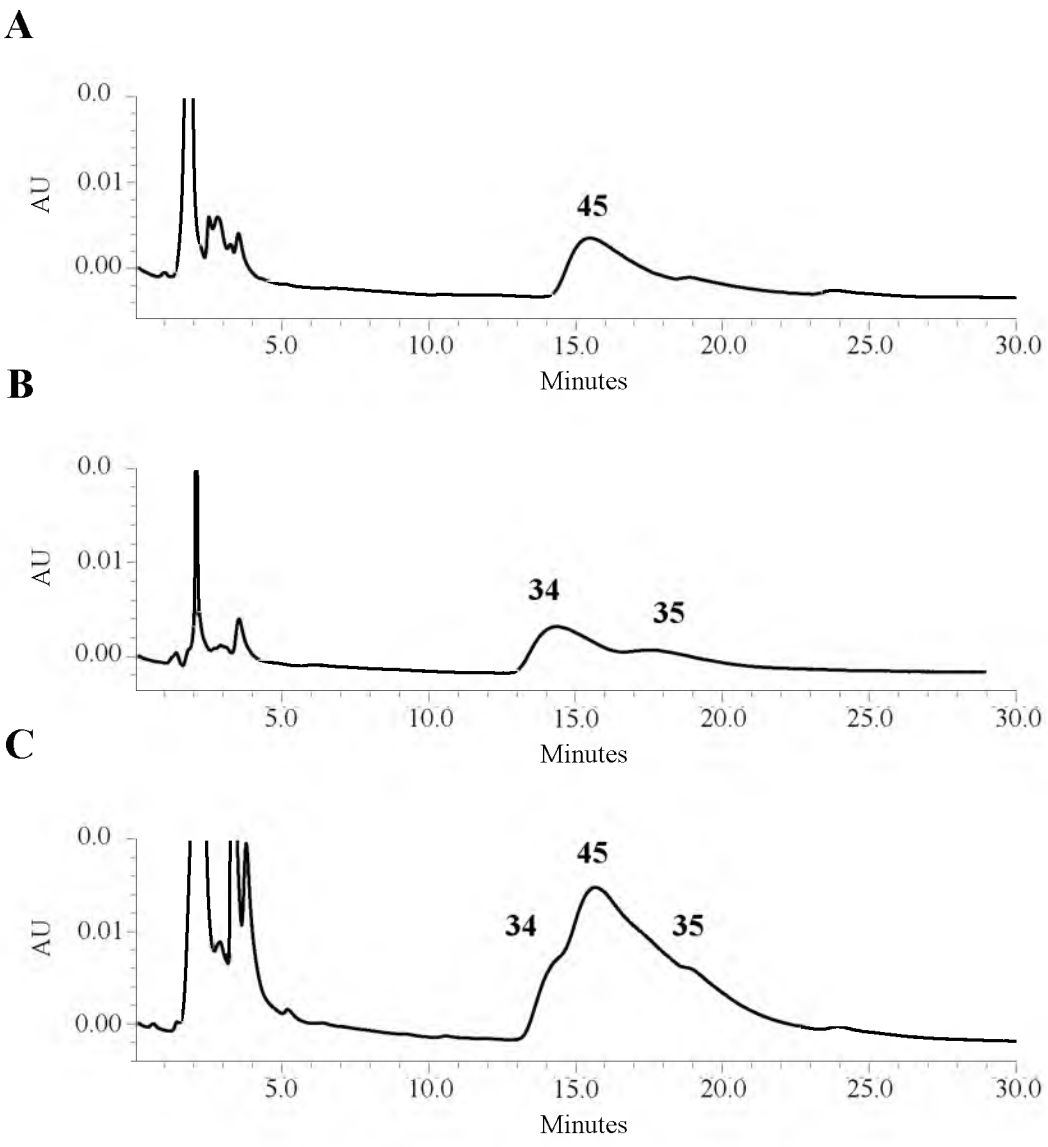


Figure 4.5. HPLC chromatograms of prenylated 4-methyltryptophan product co-injection. A, compound 45; B, compounds 34 and 35 standard; C, co-injection of compounds 45, 34, and 35. All traces were monitored at 225 nm.

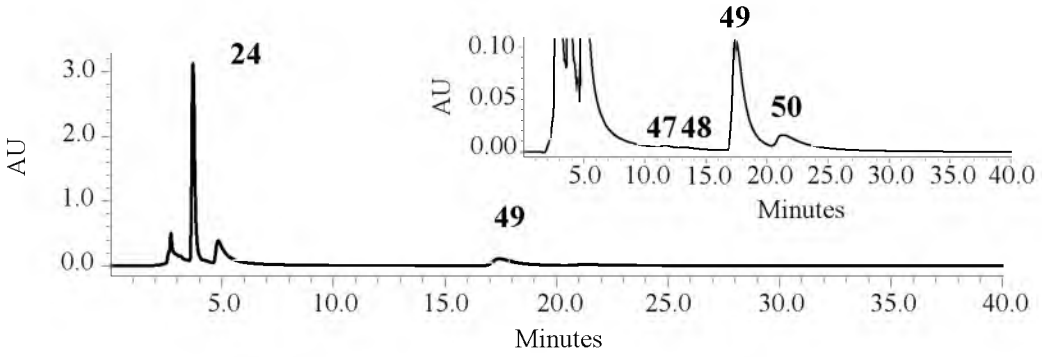


Figure 4.6. HPLC chromatogram of 4-methoxy-L-tryptophan and DMAPP incubation with 4-DMATS. Product formation was monitored at 254 nm. Inset, identical trace with adjusted y-axis for product peak visualization.

and reverse prenylated at C3 and **50** is a normal prenylated indole, most likely at N1 or C7 (see next section).

4-Aminotryptophan (21)

As described for compounds **24** and **25**, incubation of **21** with 4-DMATS gave products **51** (10.2 min, 82%) and **52** (14.2 min, 18%) (Figure 4.7). HRMS analysis indicated monoprenylation of 4-aminotryptophan. The UV spectra for **51** and **52** suggested that the indole nucleus was intact.

NMR analysis (Table 4.4) revealed the dimethylallyl moiety in compound **51** was attached in a normal orientation as evidenced by the distinctive signals for the methylene protons at C1' (3.24 ppm, d, J = 7.2), the proton at C2' (5.19 ppm, t, J = 7.2 Hz), and the methyl groups at 1.66 and 1.61 ppm. Signals for the proton at C2 (7.02, s) and two other aromatic protons ((6.89 ppm, d, J = 8.4 Hz) and (6.87 ppm, d, J = 8.4 Hz)) indicated that prenylation had occurred at either the 5 or 7 position of the indole ring. The ROESY spectrum showed correlations between the methylene protons of the prenyl moiety and the aromatic proton at C6 (6.87 ppm). The HMBC spectrum showed correlations between the methylene protons of the prenyl moiety and C3a (120.0 ppm), C4 (138.4 ppm), and C6 (127.0 ppm). Collectively, these data confirm that **51** is formed by a normal prenylation at the C5 position of 4-aminotryptophan.

Compound **52** also contains normal prenylation as confirmed by the presence of proton signals for methylene protons at C1' (3.35, d, J = 7.8), the proton at C2' (5.31 ppm, t, J = 7.8), and the methyl groups at 1.62 and 1.61 ppm. Similar to compound **51**, signals for the proton at C2 (7.05, s) and two other aromatic protons ((6.75 ppm, d, J = 7.2 Hz) and (6.36 ppm, d, J = 7.2 Hz)) indicated that prenylation had occurred at either

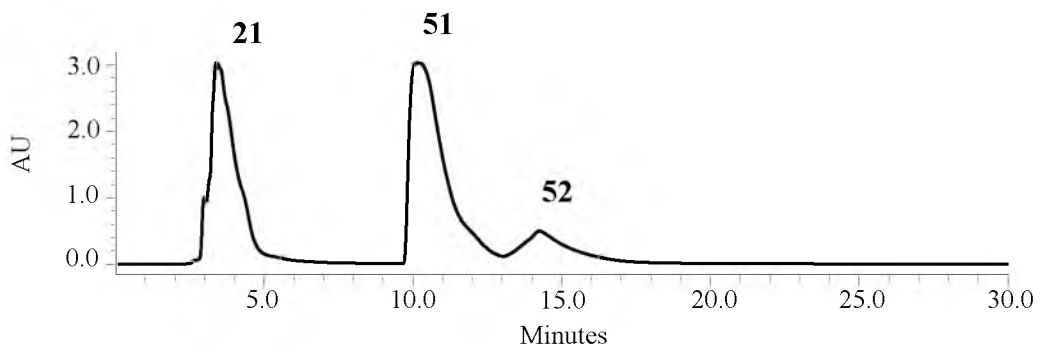
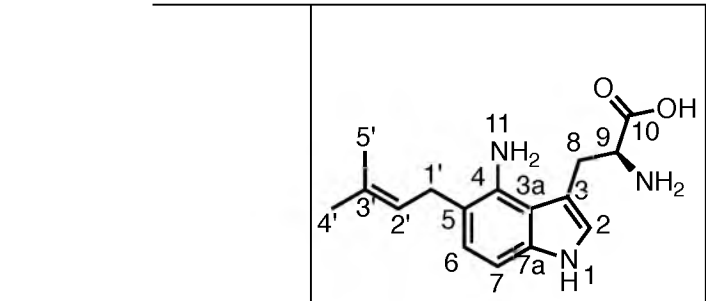
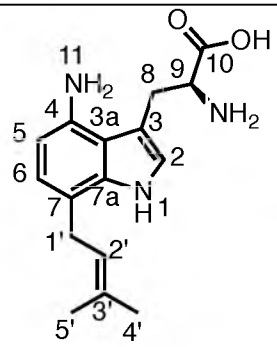
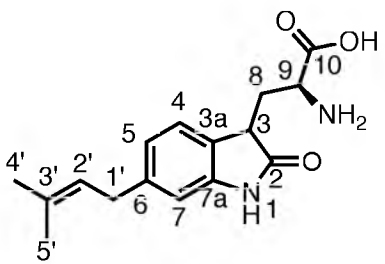


Figure 4.7. HPLC chromatogram of 4-amino-L-tryptophan and DMAPP incubation with 4-DMATS. Product formation was monitored at 225 nm.

Table 4.4. ^{1}H and ^{13}C NMR data and structures of isolated enzymatic products from 4-amino and 2-oxo-L-tryptophan

Structure/ Compound	 51		 52		 56	
Position	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**	**δ_H, multi., *J***	**δ_C**
1	exc.	-	exc.	-	exc.	-
2	7.02, s	127.3	7.05, s	126.6	-	n.f.
3	-	109.9	-	111.3	exc.	n.f.
3a	-	120.0	-	119.9	-	120.8
4	-	138.4	-	139.9	6.71, d, 8.4	121.2
5	-	n.f.	6.36, d, 7.2	110.1	7.16, d, 8.4	138.5
6	6.87, d, 8.4	127.0	6.75, d, 7.2	124.1	-	133.1
7	6.89, d, 8.4	107.0	-	120.3	7.52, s	132.5
7a	-	139.4	-	138.9	-	150.9
8	3.42-3.47, m	31.0	3.34-3.40, m	31.4	3.30, dd, 6.6, 11.4	34.6
	3.17, dd, 8.4, 15.6	-	3.15, dd, 7.2, 15.0	-	3.42-3.47, m	-
9	3.77-3.82, m	58.7	3.68-3.72, m	58.9	3.87, dd, 4.2, 6.6	53.9
10	-	n.f.	-	n.f.	-	n.f.
11	exc.	-	exc.	-	-	-
1'	3.24, d, 7.2	31.9	3.35, d, 7.8	31.2	3.16, d, 7.8	35.1
2'	5.19, t, 7.2	125.0	5.31, t, 7.8	124.2	5.26, t, 7.8	125.8
3'	-	136.8	-	137.3	-	136.8
4'	1.66, s	19.6	1.62, s	19.4	1.61, s	19.5
5'	1.61, s	27.4	1.61, s	27.3	1.61, s	27.4

δ_H and δ_C in ppm; *J* in Hz; n.f. = not found; exc. = exchanged

the 5 or 7 position of the indole ring. The COSY spectrum showed correlations between the methylene protons of the prenyl moiety and the aromatic proton at C6 (6.75 ppm). The HMBC spectrum showed correlations between the methylene protons of the prenyl moiety and C7 (120.3 ppm) and C7a (138.9 ppm), as well as a correlation between the aromatic proton at C6 and C1' (31.2 ppm). Collectively, these data confirm that **52** is formed by a normal prenylation at the C7 position of 4-aminotryptophan. Interestingly, the amino group in 4-aminotryptophan is not alkylated. Thus, 4-DMATS catalyzes the normal prenylation at C5 and C7 of 4-aminotryptophan. These experiments demonstrate that 4-DMATS catalyzes normal or reverse prenylation of tryptophan and C4-substituted analogue at five distinct sites on the indole ring, N1, C3, C4, C5, and C7, depending on the substituent (Figure 4.8).

The indole ring in tryptophan is susceptible to electrophilic alkylation at N1, at C2, C5, and C7 by direct conjugation with N1 through the benzene ring, and at C3, C4, and C6 by conjugation with N1 through the C2-C3 double bond. When C4 is blocked, 4-DMATS catalyzes alkylation at all of the activated positions, except weakly activated C6, depending on the electron-donating properties of the blocking substituent. In addition, both normal and reverse prenylation is seen at C3 for 4-methyltryptophan. Alkylation of the indole ring at sites separated by several angstroms and by normal and reverse alkylation of C3 indicates that 4-DMATS can bind DMAPP in different conformations. The high regioselectivity for C4 alkylation with the normal substrates suggests that binding and catalysis of a single conformation has been optimized for synthesis of 4-dimethylallyltryptophan, but when alkylation at C4 is blocked, other conformations become competitive.

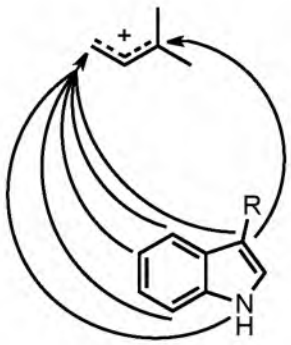


Figure 4.8. Prenylation sites and orientations catalyzed by 4-DMATS.

The methyl, methoxy, and amino substituents introduced at C4 are all electron-donating relative to hydrogen. For the methyl substituent ($\sigma^+ = 0.31$),[144] alkylation is mostly at C3, giving normal and reverse prenylated indolines, along with a smaller amount of the N1 alkylated indole. The methoxy group ($\sigma^+ = 0.78$) directs alkylation predominantly to C5 giving only trace amounts of putative indolines. The amino group ($\sigma^+ = 1.3$) directs alkylations exclusively to C5 and C7. Thus, when C4 is blocked by a relatively weak electron donating substituent, the dimethylallyl cation alkylates C3 to generate normal and reverse prenylated iminium intermediates, which cyclize to the corresponding indolines. The more powerful electron-donating methoxy and amino groups direct alkylation to the *ortho* (C5) and *para* (C7) carbons, positions further activated by N1. Interestingly, the amino group in 4-aminotryptophan is not alkylated.

4-Hydroxy-L-tryptophan (22) and Dihydroiso-L-tryptophan (29)

In addition to electronic effects, the size differences between a hydrogen and methyl, methoxy, and amino substituents may cause steric clashing between the aromatic amino acid and active site residues causing reorientation of the active site molecules. The

prenylation pattern of 4-hydroxy-L-tryptophan (**22**) by 4-DMATS, an electron-donating yet smaller substrate than **24**, would help confirm the dominating factor.

Incubation of **22** with DMAPP and 4-DMATS yielded product **53** (13.5 min) (Figure 4.9). UV analysis of **53** (λ_{max} = 222.7, 259.4, 287.9, 295.0 nm) indicated the indole ring was still intact. During overnight reactions and purification of compound **53**, it was evident the product was oxidizing to an unknown product. Attempts to prevent product oxidation (reactions performed under nitrogen/argon) and thus **53** loss were unsuccessful; therefore analysis by NMR spectroscopy and MS was not performed. It is hypothesized that prenylation does not occur on the substituent oxygen, but rather a position on the indole ring such as C5 or C7, as is seen with 4-aminotryptophan. However, the lack of two distinct products raises the question as to whether **22** is prenylated at the C5 or C7 position of the indole ring.

As stated previously (Chapter 3), DIT (**29**) provides an interesting substrate analogue as it contains an indoline ring with an amino acid moiety attached on the ring's nitrogen atom. As with SirD, because the amino acid moiety is important for catalysis the connectivity of the bicyclic ring to the amino acid moiety may allow new sites to be

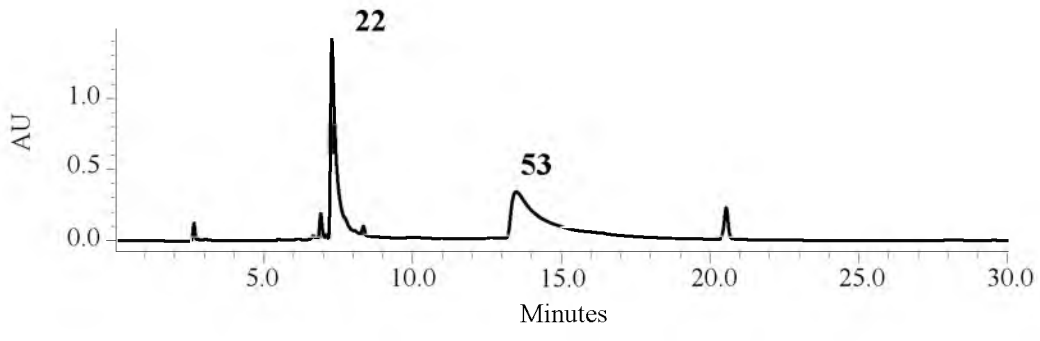


Figure 4.9. HPLC chromatogram of 4-hydroxy-L-tryptophan and DMAPP incubation with 4-DMATS. Product formation was monitored at 225 nm.

positioned for prenylation. The 2,3-dihydropyrrole ring with a tertiary amine also prevents prenylation from occurring at the N1, C2 and C3 positions of the indoline ring. The disruption of the pyrrole ring's aromaticity also eliminates the activation of the C4 and C6 carbons.

Incubation of 4-DMATS with compound **29** and DMAPP gave products **54** (11.1 min, 36%) and **55** (26.4 min, 64%) (Figure 4.10). Similar to the SirD reaction with **29,** UV analysis of **54** indicated the indoline ring of DIT had changed (λ_{max} = 220.1, 268.9 nm). UV analysis of the product **55** revealed the indoline ring of DIT had not changed (λ_{max} = 205.0, 249.9, 298.6 nm), signifying prenylation occurred on the benzene ring of DIT and no other modifications had occurred at the five-membered 2,3-dihydropyrrole ring. Although the prenylation sites are undetermined, the ability of 4-DMATS to prenylate DIT illustrates its exceptional promiscuity.

2-Oxo-L-tryptophan

Recently, Luk and coworkers reported that the K174A mutant of *A. fumigatus* 4-DMATS catalyzes reverse prenylation of C3.[89] A crystal structure of FgaPT2 and kinetic studies indicate that K174 is an important base, which facilitates removal of the C4 hydrogen during the electrophilic aromatic substitution reaction.[59] They suggested the reaction catalyzed by the wild-type enzyme proceeds in two steps, reverse alkylation at C3 followed by a Cope rearrangement and deprotonation to form the normal C4 prenylation product.[89] To investigate this claim, two diastereomers of reverse C3 prenyl-2-oxo-L-tryptophan (**0393-1** and **0393-2**) were incubated with wild type 4-DMATS to determine whether cyclization of the dimethylallyl moiety occurs. Unfortunately, only the

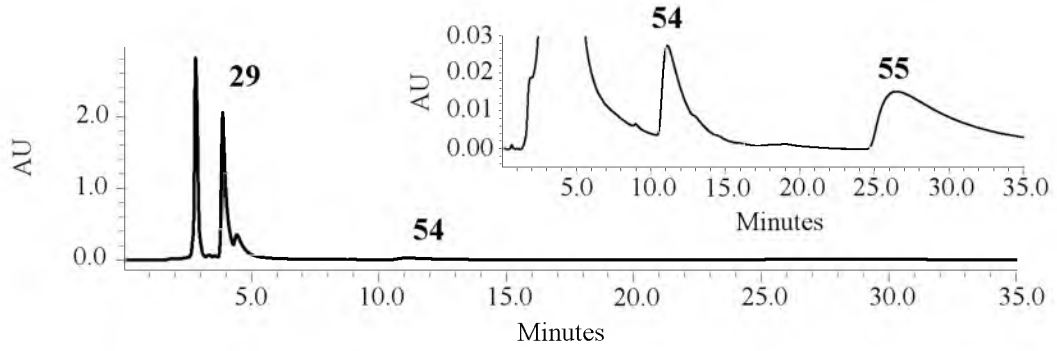


Figure 4.10. HPLC chromatogram of dihydroiso-L-tryptophan and DMAPP incubation with 4-DMATS. Product formation was monitored at 225 nm. Inset, identical trace with adjusted y-axis for product peak visualization.

2-oxo derivatives were available for use as the reverse C3 prenyl-L-tryptophan derivatives were unstable.[147]

Compounds **0393-1** and **0393-2** were individually incubated overnight in the presence and absence of 4-DMATS. HPLC analysis indicated no changes to the elution profile occurred when enzyme was introduced to the incubation (Figure 4.11). Furthermore, incubations with 4-DMATS and DMAPP or inorganic pyrophosphate yielded no changes to the chromatograms (data not shown). These results prove that 4-DMATS neither rearranges nor prenylates **0393-1** or **0393-2**. The 2-oxo modification of the tryptophan side chain may prevent 4-DMATS from binding the modified substrate. To test 2-oxotryptophan as an alternate substrate, 2-oxo-L-tryptophan was incubated with DMAPP and 4-DMATS. One product (**56**, 10.5 min) was formed (Figure 4.12) and UV analysis revealed a spectrum (λ_{max} = 229.8, 255.8, 373.8 nm) similar to that of the substrate (λ_{max} = 223.8, 254.6, 362.2 nm).

Analysis by MS was inconclusive, however prenylation was unmistakable when the ^{1}H NMR spectrum (Table 4.4) was examined. The dimethylallyl moiety in **56** was

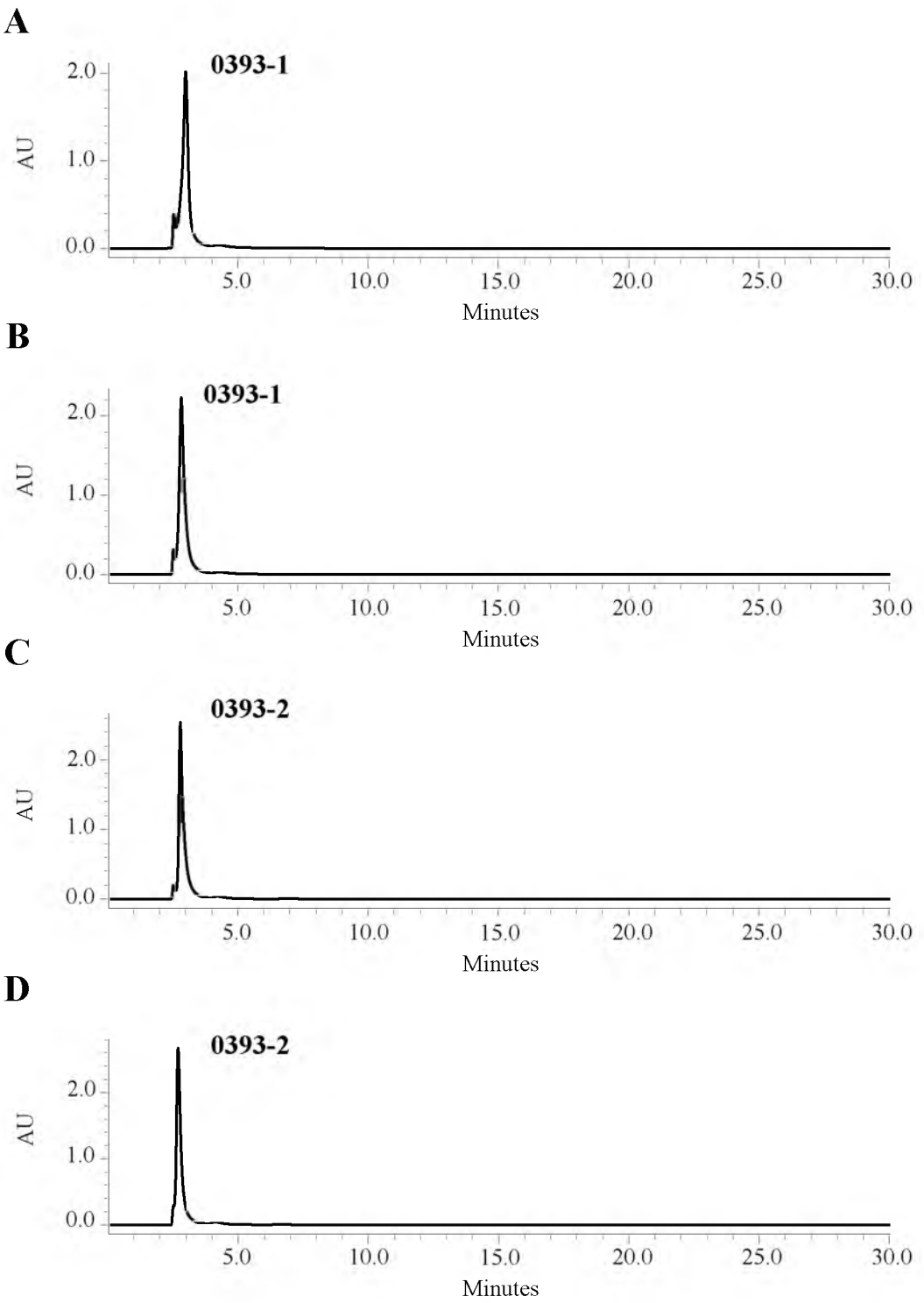


Figure 4.11. HPLC chromatograms of prenylated 2-oxo-L-tryptophan incubations with 4-DMATS. A, compound 0393-1 control; B, compound 0393-1 incubated with 4-DMATS; C, compound 0393-2 control; D, compound 0393-2 incubated with 4-DMATS. All traces were monitored at 225 nm.

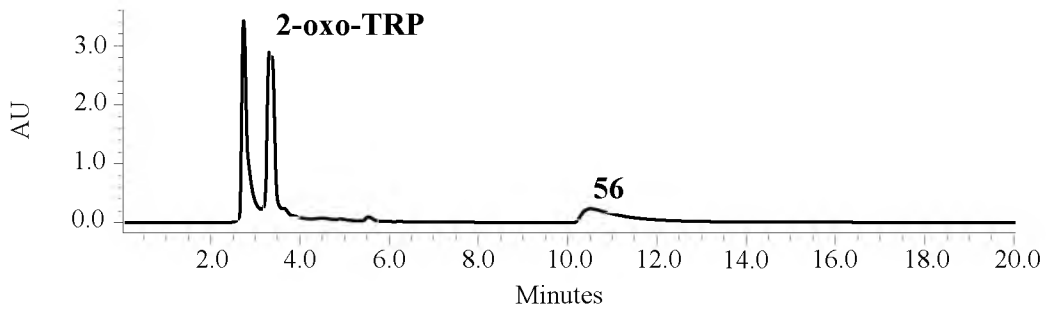


Figure 4.12. HPLC chromatogram of 2-oxo-L-tryptophan and DMAPP incubation with 4-DMATS. Product formation was monitored at 225 nm.

attached in a normal orientation as evidenced by the distinctive signals for the methylene protons at C1' (3.16 ppm, d, J = 7.8), the proton at C2' (5.26 ppm, t, J = 7.8 Hz), and two methyl groups at 1.61 ppm. Signals for three aromatic protons, one a downfield singlet (7.52, s) and two coupled doublets ((7.16 ppm, d, J = 7.8 Hz) and (6.71 ppm, d, J = 7.8 Hz)) indicated that prenylation had occurred at either the 5 or 6 position of the indole ring. Therefore, an aromatic singlet must be placed at C4 or C7. Since the proton at C7 is in close proximity to the ring nitrogen, its deshielded nature should push the chemical shift downfield. This suggests prenylation occurred at the C6 position. Analysis of compound **56** by 2D NMR spectroscopy was ambiguous when determining prenylation site. Together, 4-DMATS' ability to accept 2-oxotryptophan and catalyze prenylation at a position other than C4 with its inability to rearrange the reverse C3 prenyl-2-oxo-L-tryptophan derivatives supports a direct electrophilic aromatic substitution mechanism. The lack of rearrangement from C3 to C4, and in view of the multiplicity of products we observe when C4 is blocked, including normal and reverse alkylation at C3, a Cope rearrangement to move the dimethylallyl moiety from C3 to C4 seems to be an

unnecessary embellishment. Most likely, alkylation at C3 by the K174A mutant is a result of a decrease in the rate of alkylation at C4, which allows electrophilic attack at C3 to compete.

Overall, our results provide strong evidence for a dissociative electrophilic alkylation of the indole nucleus by C1', and in one instance C3', of the dimethylallyl cation generated by heterolysis of DMAPP. 4-Dimethylallyltryptophan is the only product of the normal reaction catalyzed by 4-DMATS. When alkylation at C4 is blocked, the enzyme no longer catalyzes a regioselective reaction, but instead produces a variety of products characteristic of alkylation of the nucleophilic positions (N1, C3, C5, and C7) on the indole ring and electrophilic attack by the dimethylallyl cation. In addition, we saw only a three-fold variation in k_{cat} for the methyl, methoxy, and amino tryptophans, which is most readily explained by formation of the dimethylallyl cation as the rate-limiting step, followed by alkylation of the indole nucleus. These results are consistent with previous studies with 7-methyl and 7-methoxy derivatives of tryptophan. In this case, tryptophan and the 7-substituted analogues are alkylated at C4 with only a two-fold variation in k_{cat}.[80]

Conclusion

The work presented in this dissertation describes the synthesis of aromatic amino acid analogues, the biochemical characterization of a new aromatic amino acid prenyltransferase SirD, and the promiscuous nature and multisite prenylation capabilities of SirD and 4-DMATS.

Two tyrosine/phenylalanine and ten tryptophan derivatives were synthesized using a combination of synthetic and biosynthetic methods. Notably, tryptophan

synthase was utilized to synthesize enantiomerically pure L-tryptophan analogues from a variety of substituted indoles. The dimethylallyltyrosine synthase SirD was cloned, heterologously expressed, and biochemically characterized with its natural substrates L-tyrosine and DMAPP. Using a secondary structure predictor and homology modeler, SirD was presented as the newest member of the αββα aromatic prenyltransferase family. SirD and the dimethylallyltryptophan synthase 4-DMATS were tested for substrate promiscuity. SirD was found to catalyze *O*-, *N*-, and *S*-prenylations of tyrosine derivatives. SirD was also determined to catalyze normal prenylations at the C6 and C7 positions and reverse prenylations at the N1 and C7 positions of the indole ring of tryptophan derivatives. 4-DMATS was found to prenylate 4-substituted tryptophans at four different positions on the indole ring with normal or reverse prenylation at one of the sites. Normal prenylation occurred at positions N1, C3, C5 and C7 while reverse prenylation occurred at C3. With natural prenylation occurring at C4, 4-DMATS was shown to prenylate five of the seven sites for alkylation of the indole nucleus by the DMATS superfamily, providing logical rationale for the evolution of members of the superfamily from a common ancestor.

In conclusion, the ability of the aromatic amino acid prenyltransferases SirD and 4-DMATS to catalyze both normal and reverse prenylations on six of the seven available positions of the indole ring (C2 the exception) supports a dissociative electrophilic alkylation mechanism where orientation of the substrates within the active site and substituent electronic effects determine the position and type of prenylation. Although many substrates, reactions, and products were studied, much is left unfinished. Several analogues were shown to be substrates yet their products were not determined.

Establishing these prenylation products will undoubtedly increase our understanding of prenyltransferase promiscuity. Understandably, crystal structures with complexed substituted tryptophans and/or substrate docking studies would also be valuable additions to the study of SirD and 4-DMATS, their reaction mechanisms, and their promiscuous nature.

APPENDIX A

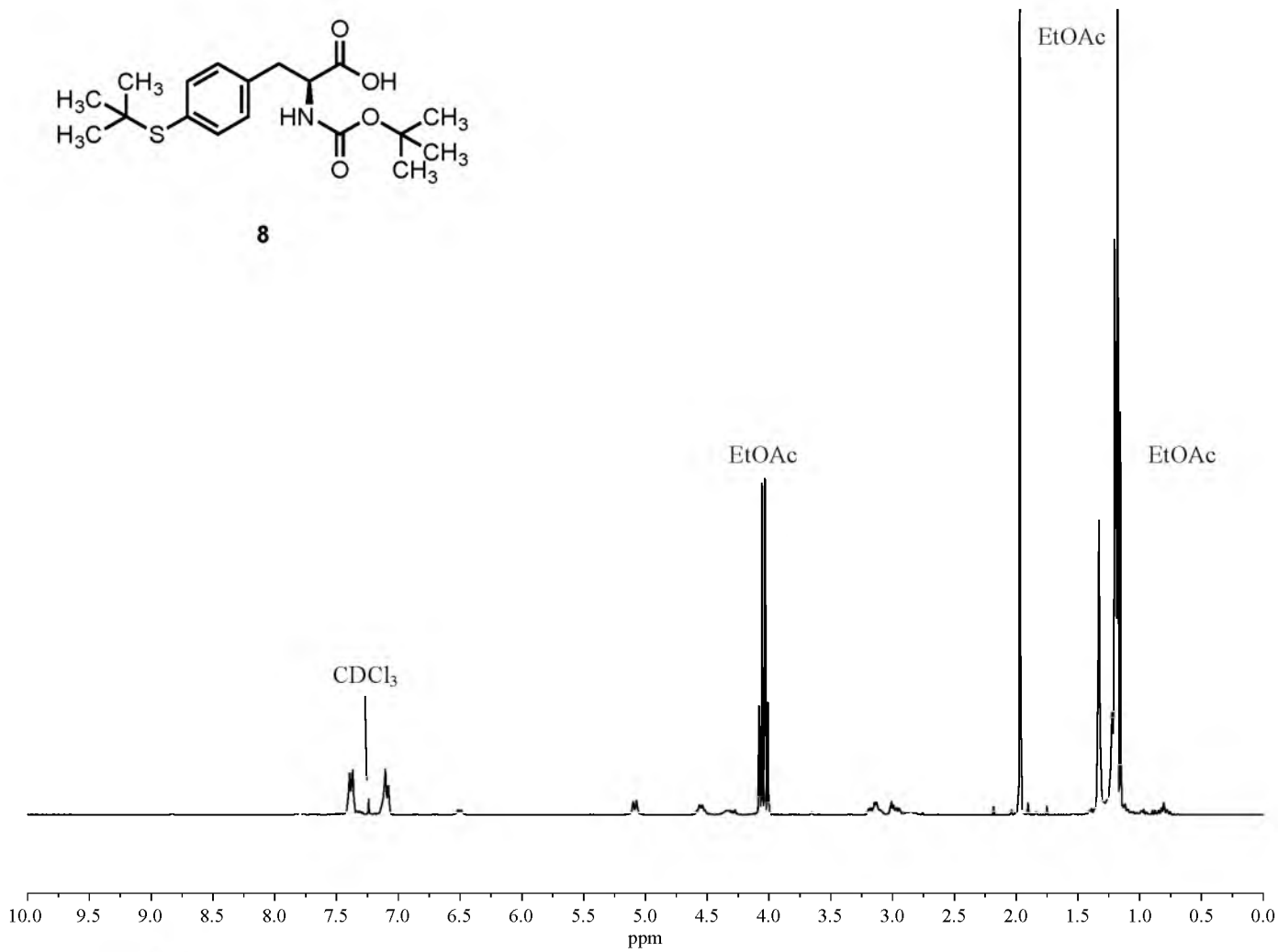
HRMS DATA

Table S.1. HRMS data of synthesized compounds

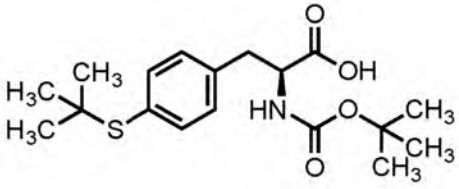
Compound	Chemical Formula	Calculated	Found	Deviation (ppm)
9	$C_{18}H_{21}N_2O_4S_2$	393.0943 $(M + H)^+$	393.0945	0.5
12	$C_{26}H_{22}NO_7NaSF_3$	572.0967 $(M + Na)^+$	572.0970	0.5
14	$C_{27}H_{25}NO_4Na$	450.1681 $(M + Na)^+$	450.1688	1.6
15	$C_{11}H_{13}NO_2Na$	214.0838 $(M + Na)^+$	214.0838	-2.8
17	$C_{12}H_{16}N_2S$	206.1003 $(M + H)^+$	206.1007	1.9
20	$C_{12}H_{14}N_2O_2Na$	241.0953 $(M + Na)^+$	241.0964	4.6
21	$C_{11}H_{14}N_3O_2$	220.1086 $(M + H)^+$	220.1078	-3.6
22	$C_{11}H_{13}N_2O_3$	221.0926 $(M + H)^+$	221.0934	3.6
23	$C_{11}H_{12}N_2O_2NaS$	259.0517 $(M + Na)^+$	259.0524	2.7
24	$C_{12}H_{14}N_2O_3Na$	257.0902 $(M + Na)^+$	257.0903	0.4
25	$C_{12}H_{15}N_2O_2$	219.1134 $(M + H)^+$	219.1136	0.9
26	$C_{13}H_{15}N_2O_2$	231.1134 $(M + H)^+$	231.1135	0.4
27	$C_{11}H_{13}N_2O_3$	221.0926 $(M + H)^+$	221.0922	-1.8
28	$C_{12}H_{14}N_2O_2Na$	241.0953 $(M + Na)^+$	241.0955	0.8
29	$C_{11}H_{15}N_2O_2$	207.1134 $(M + H)^+$	207.1148	6.8
30	$C_{14}H_{19}NO_3Na$	272.1263 $(M + Na)^+$	272.1257	-2.2
31	$C_{14}H_{19}NO_2NaS$	288.1034 $(M + Na)^+$	288.1041	2.4
32	$C_{16}H_{20}N_2O_2Na$	295.1422 $(M + Na)^+$	295.1428	2.0
33	$C_{16}H_{21}N_2O_2$	273.1603 $(M + H)^+$	273.1609	2.2
34	$C_{17}H_{23}N_2O_2$	287.1760 $(M + H)^+$	287.1762	0.7
35	$C_{17}H_{23}N_2O_2$	287.1760 $(M + H)^+$	287.1763	1.0
36	$C_{17}H_{22}N_2O_3Na$	325.1528 $(M + Na)^+$	325.1532	1.2
37	$C_{17}H_{22}N_2O_3Na$	325.1528 $(M + Na)^+$	325.1532	1.2
38	$C_{17}H_{22}N_2O_2Na$	309.1579 $(M + Na)^+$	309.1579	0.0
39	$C_{17}H_{22}N_2O_2Na$	309.1579 $(M + Na)^+$	309.1576	-1.0
42	$C_{16}H_{22}N_2O_2Na$	297.1579 $(M + Na)^+$	297.1586	2.4
43	$C_{17}H_{22}N_2O_2Na$	309.1579 $(M + Na)^+$	309.1581	0.6
44	$C_{17}H_{22}N_2O_2$	287.1760 $(M + H)^+$	287.1755	-1.7
45	$C_{17}H_{22}N_2O_2Na$	309.1579 $(M + Na)^+$	309.1594	4.9
46	$C_{17}H_{22}N_2O_2$	287.1760 $(M + H)^+$	287.1760	0.0
47	$C_{17}H_{22}N_2O_3Na$	325.1528 $(M + Na)^+$	325.1539	3.4
48	$C_{17}H_{22}N_2O_3Na$	325.1528 $(M + Na)^+$	325.1524	-1.2
49	$C_{17}H_{22}N_2O_3Na$	325.1528 $(M + Na)^+$	325.1530	0.6
50	$C_{17}H_{22}N_2O_3Na$	325.1528 $(M + Na)^+$	325.1534	1.8
51	$C_{16}H_{22}N_3O_2$	288.1712 $(M + H)^+$	288.1715	1.0
52	$C_{16}H_{22}N_3O_2$	288.1712 $(M + H)^+$	288.1717	1.7

APPENDIX B

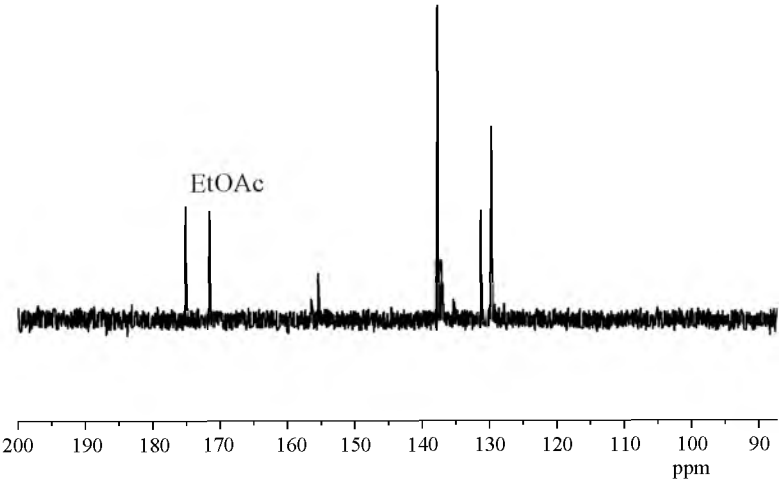
NMR SPECTRA



NMR 1. 300 MHz ^{1}H NMR spectrum of compound 8 in $CDCl_3$.

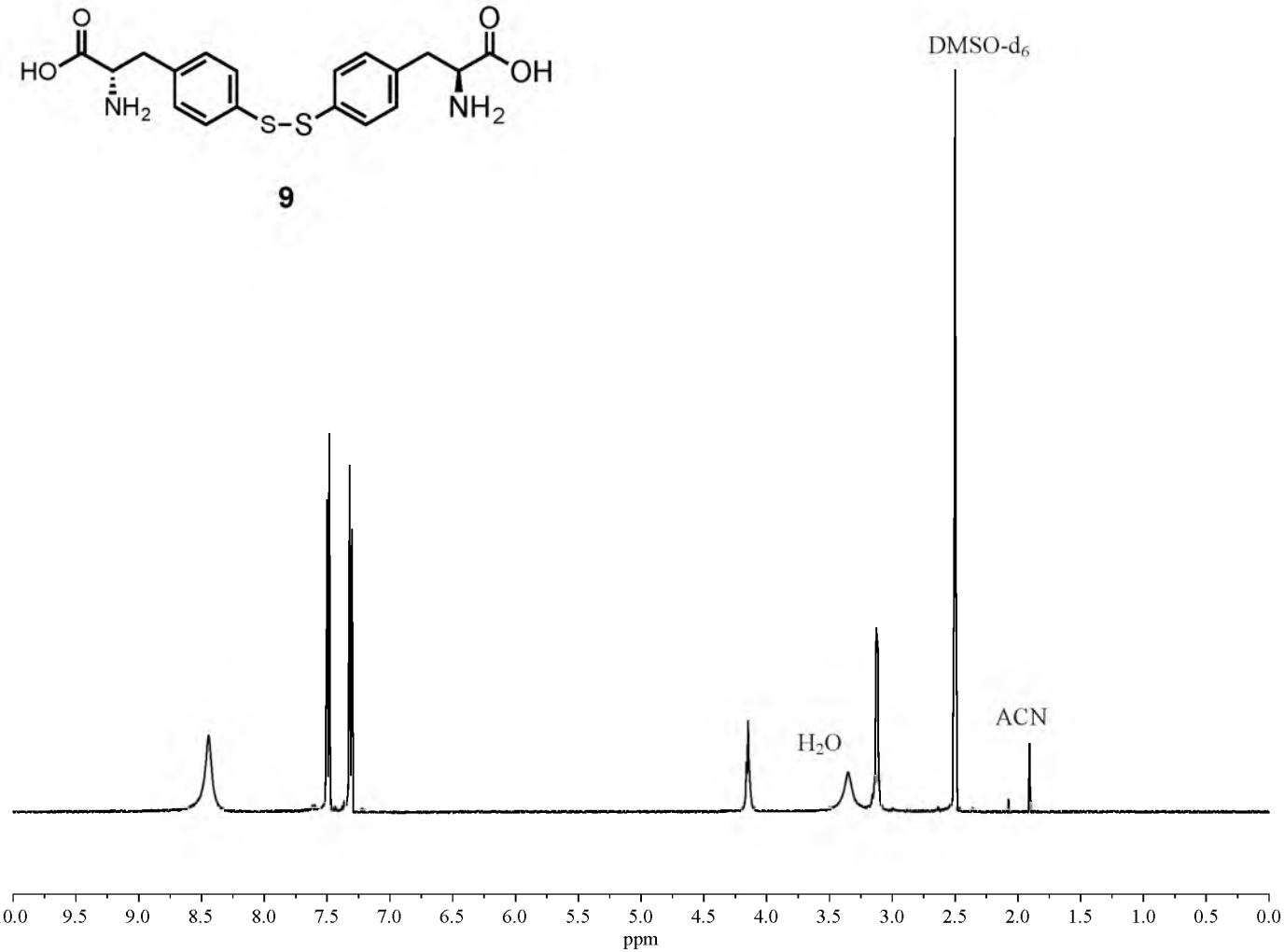


8

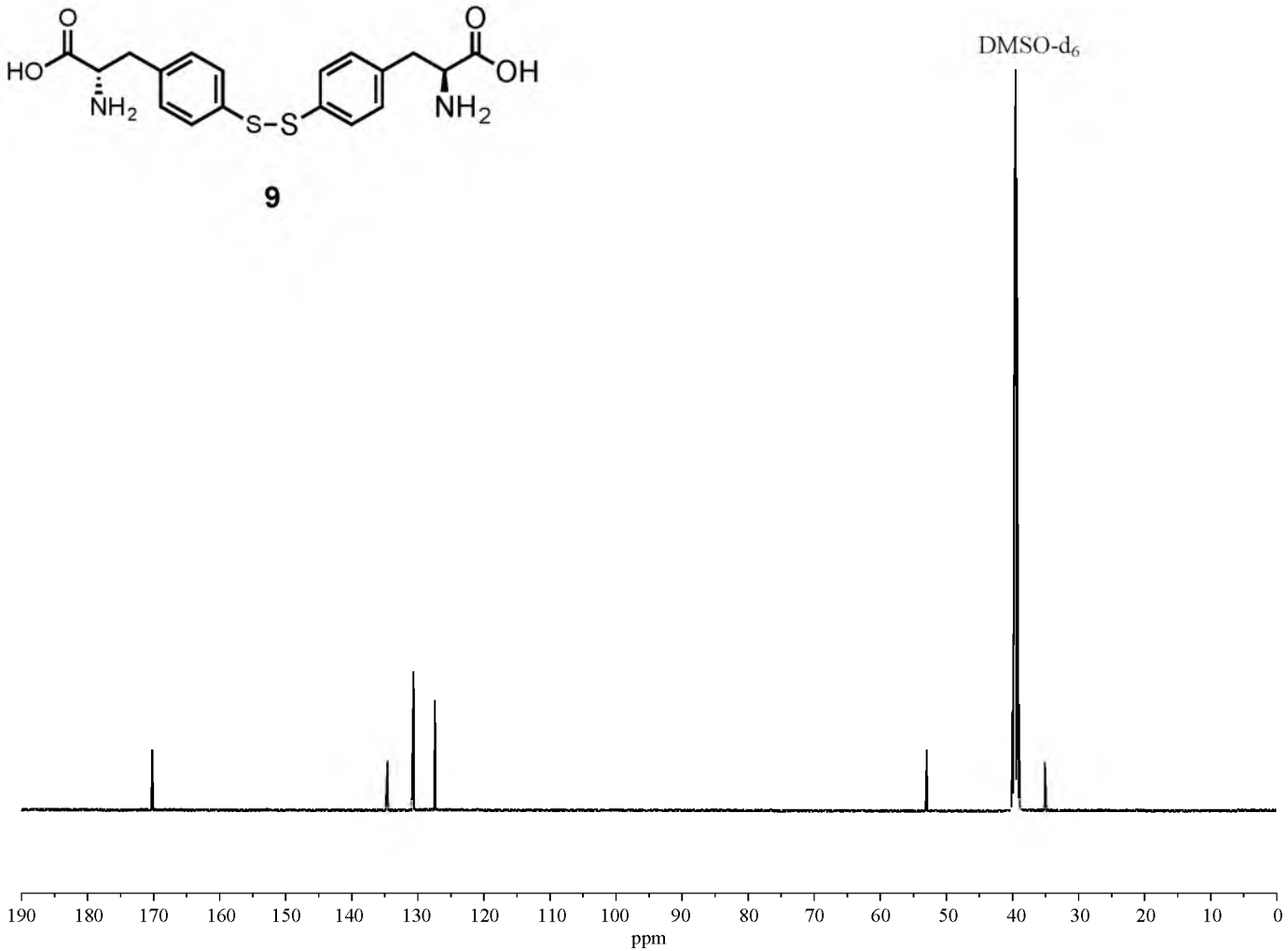


NMR 2. 75 MHz ^{13}C NMR spectrum of compound 8 in $CDCl_3$.

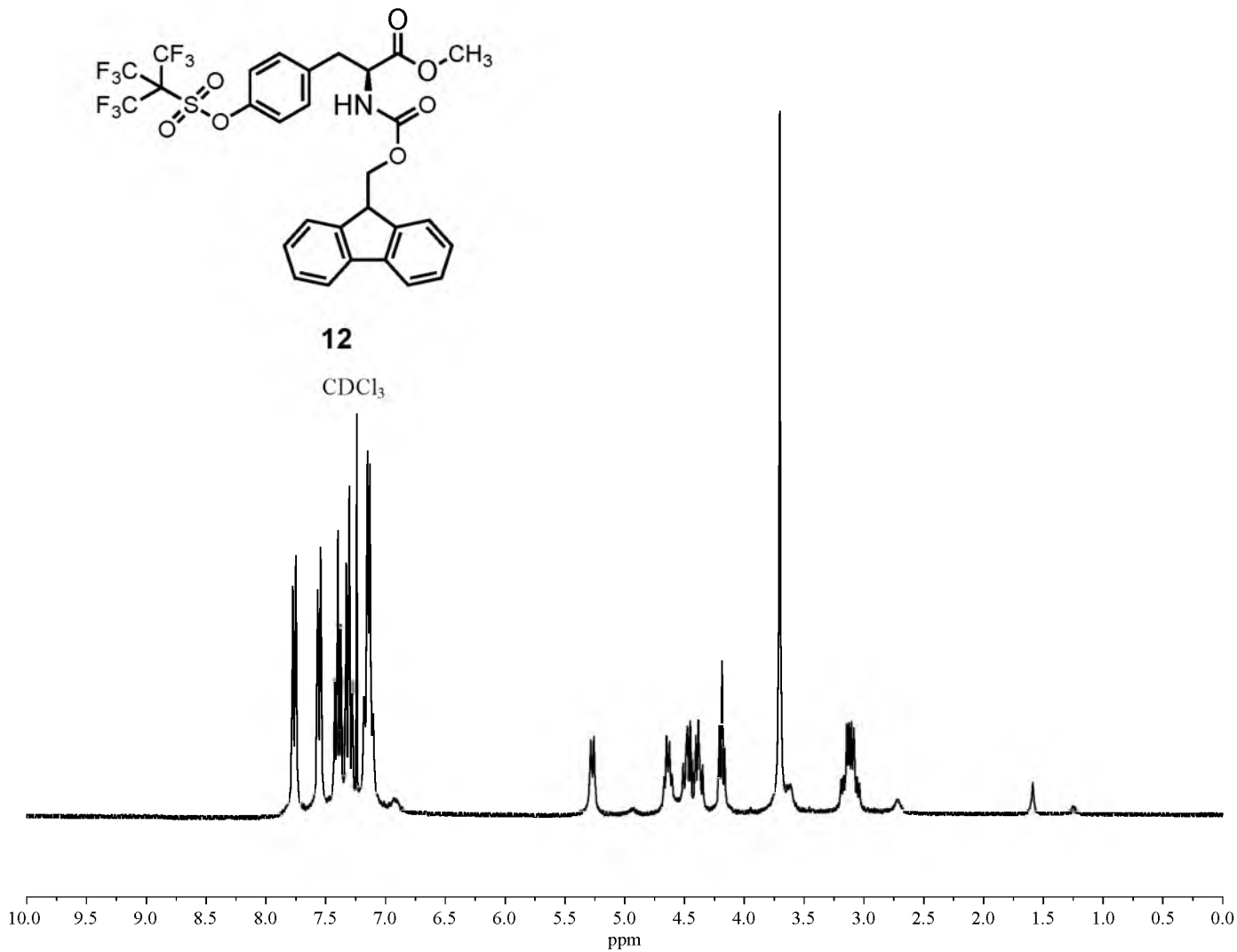
EtOAc
EtOAc
EtOAc
CDCl3
80
70
60
50
40
30
20
10
0



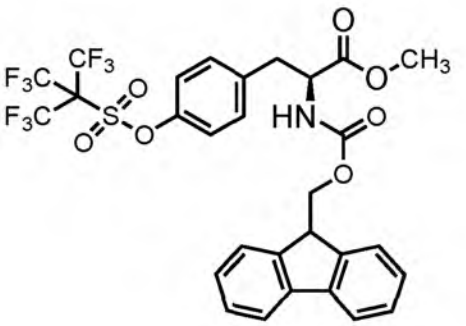
NMR 3. 500 MHz ^{1}H NMR spectrum of compound 9 in DMSO-d$_6$.



NMR 4. 125 MHz ^{13}C NMR spectrum of compound 9 in DMSO-d$_6$.

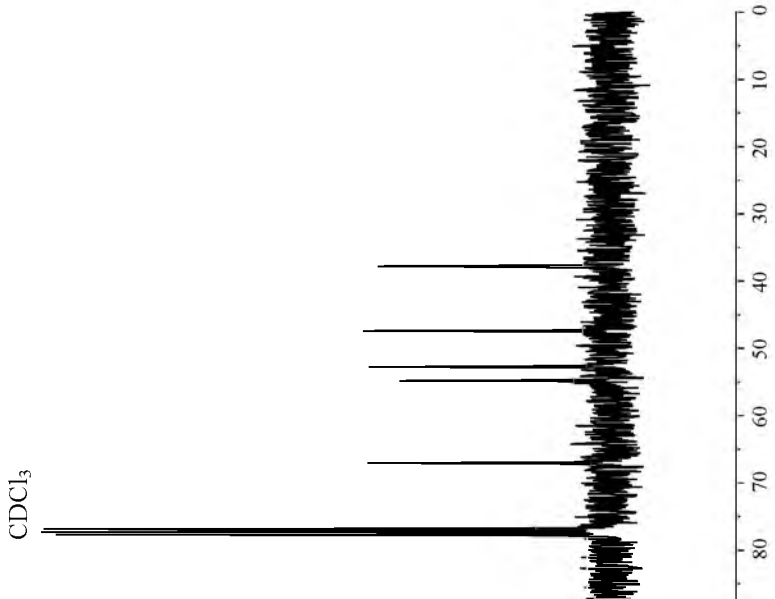


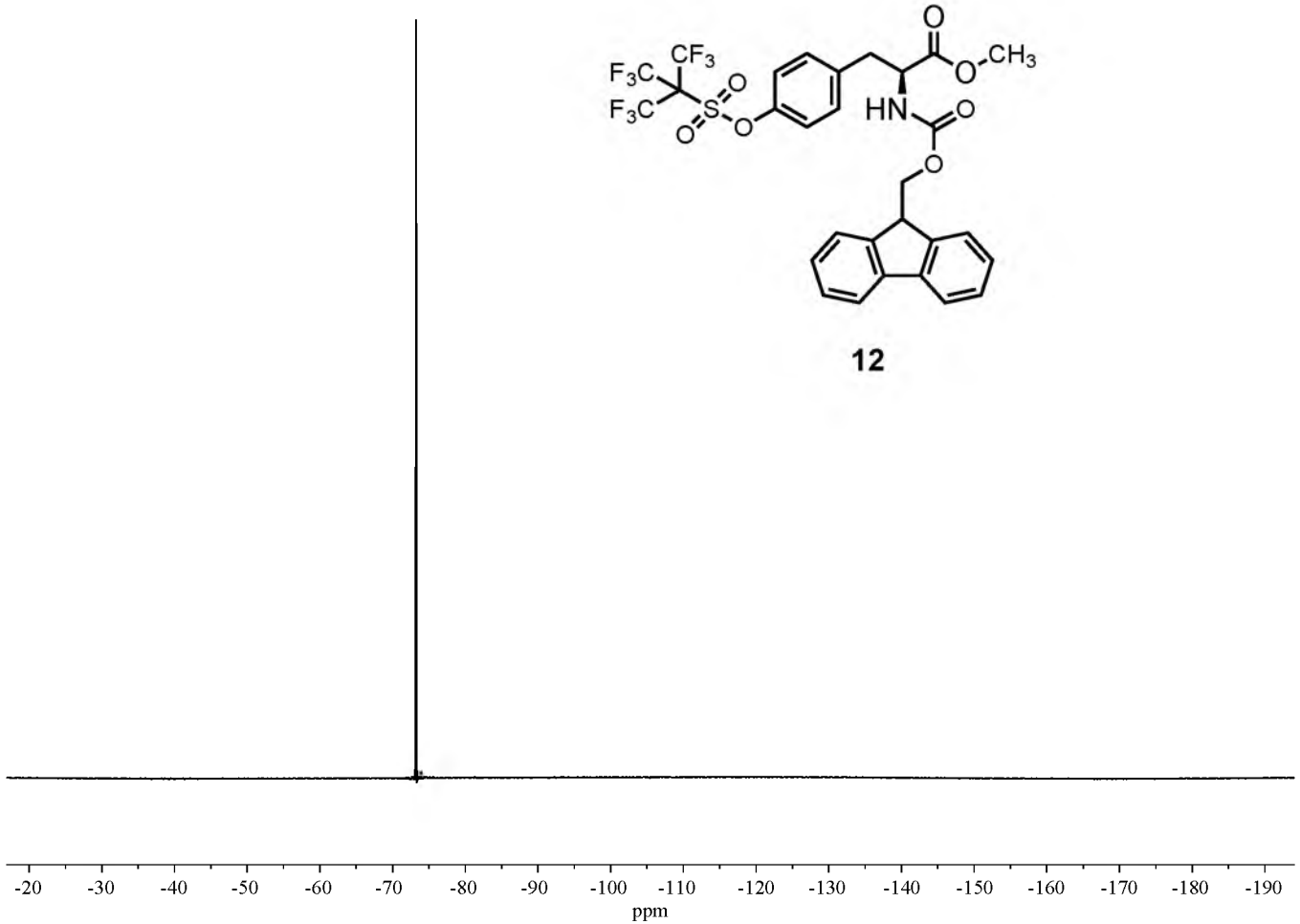
NMR 5. 300 MHz ^{1}H NMR spectrum of compound 12 in $CDCl_3$.



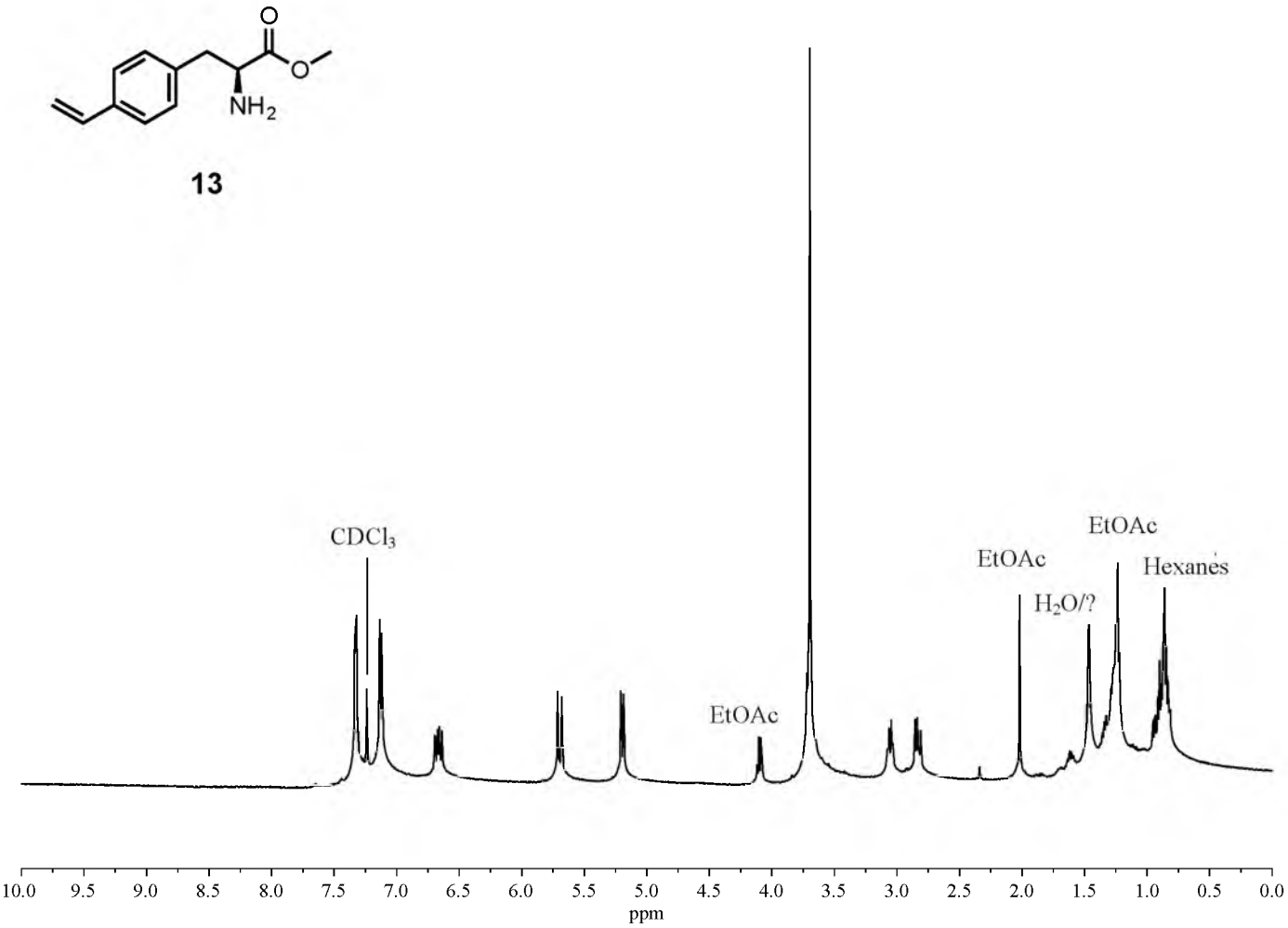
12

NMR 6. 75 MHz ^{13}C NMR spectrum of compound 12 in $CDCl_3$.

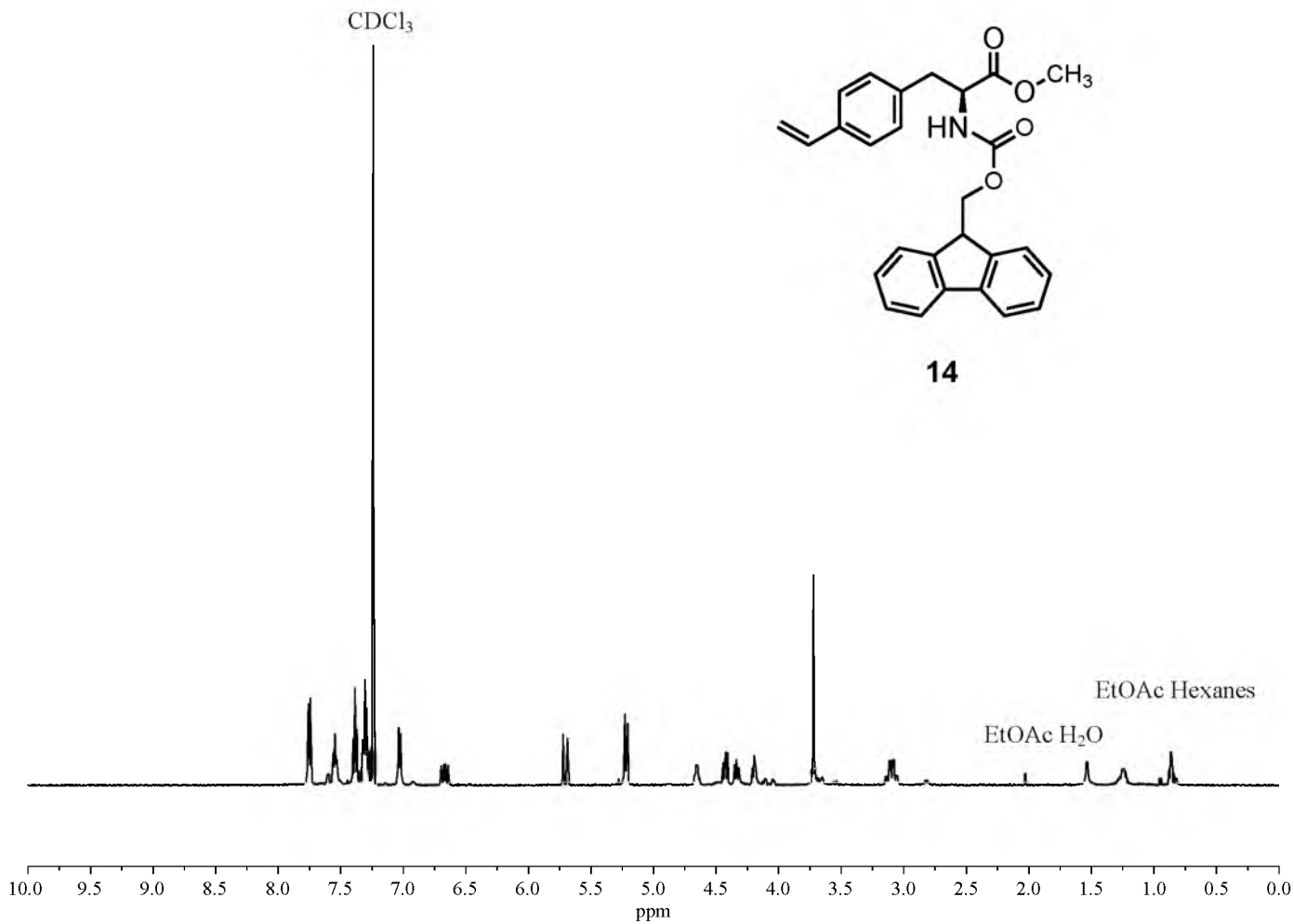

CDCl3
0
10
20
30
40
50
60
70
80



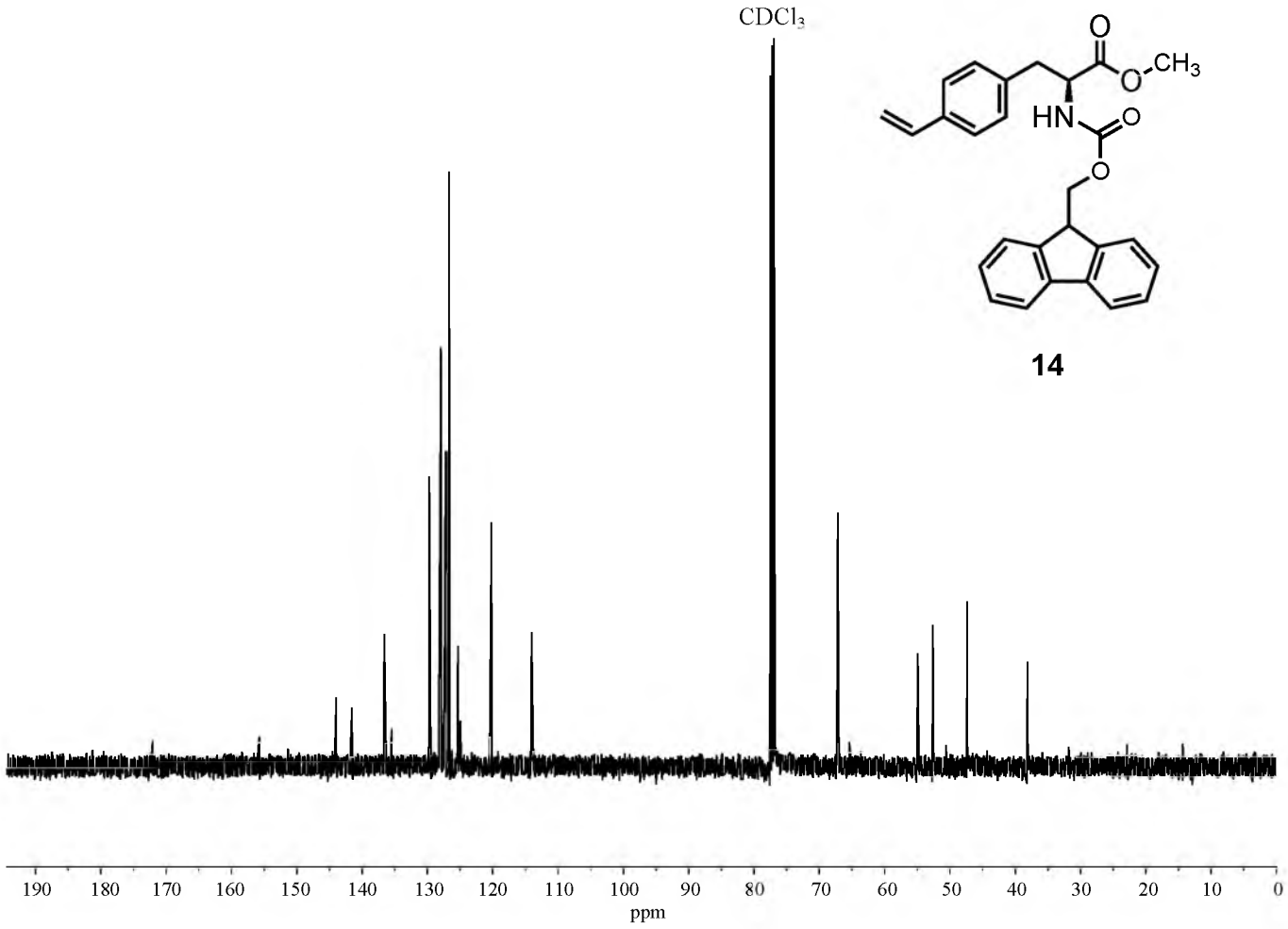
NMR 7. 282 MHz ^{19}F NMR spectrum of compound 12 in DMSO-d$_6$.



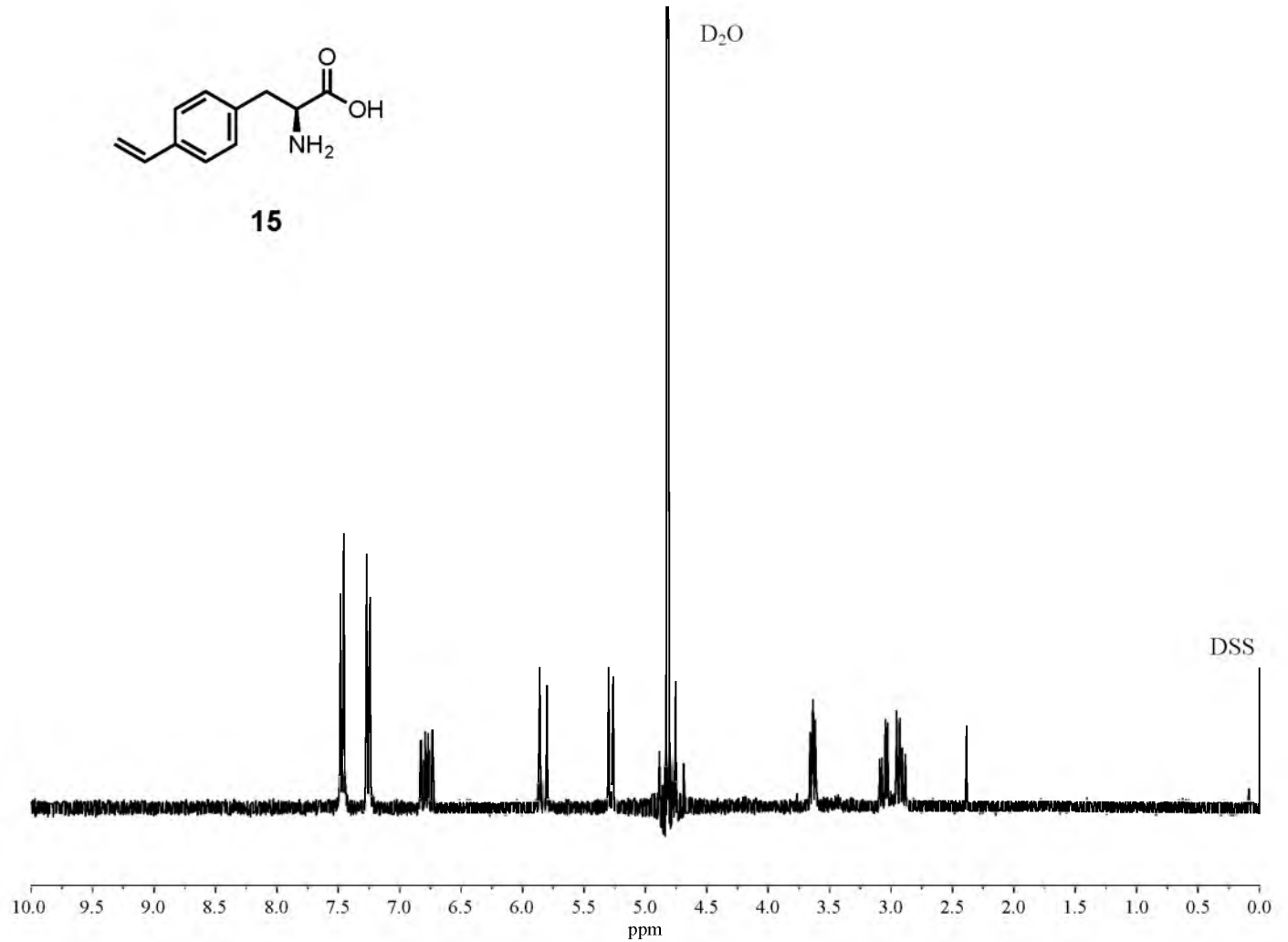
NMR 8. 500 MHz ^{1}H NMR spectrum of compound 13 in $CDCl_3$.



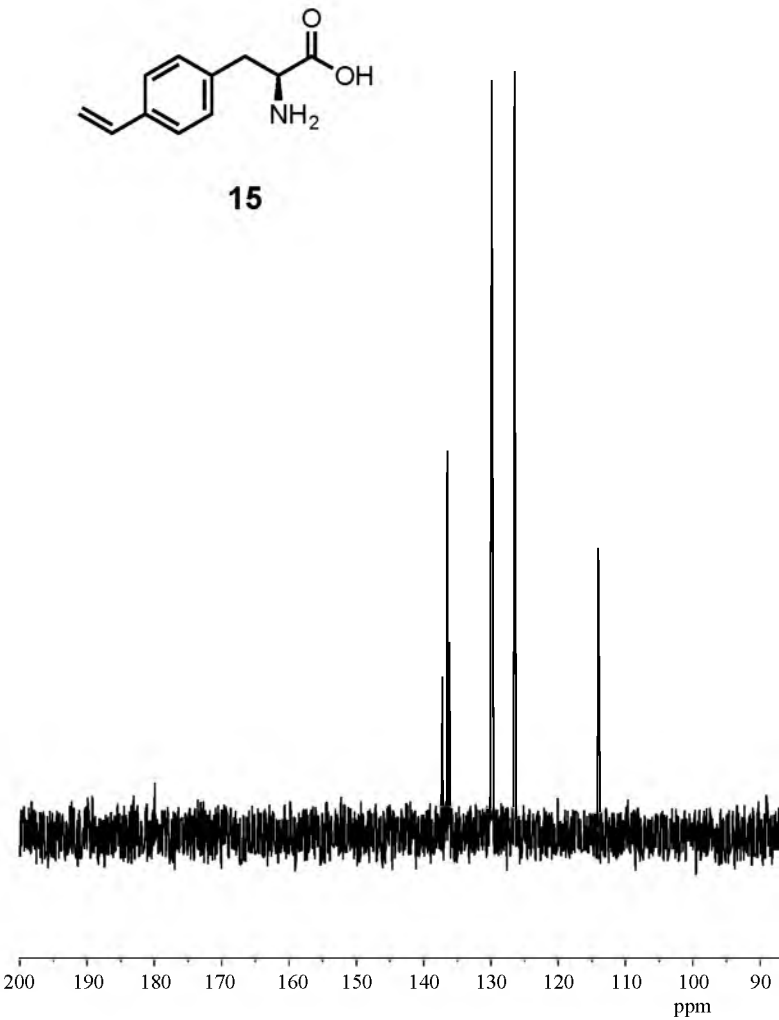
NMR 9. 500 MHz ^{1}H NMR spectrum of compound 14 in $CDCl_3$.



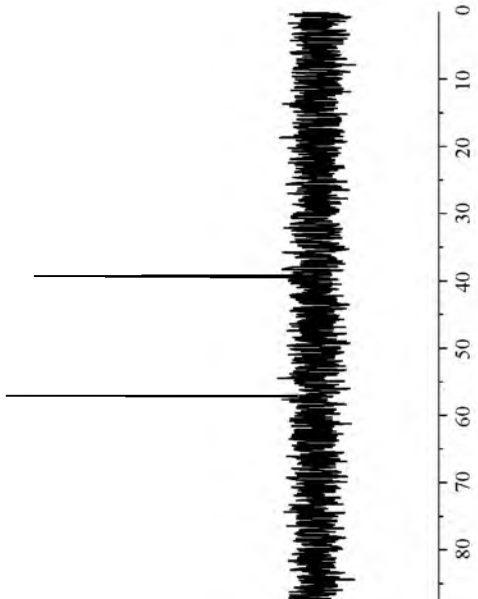
NMR 10. 125 MHz ^{13}C NMR spectrum of compound 14 in $CDCl_3$.

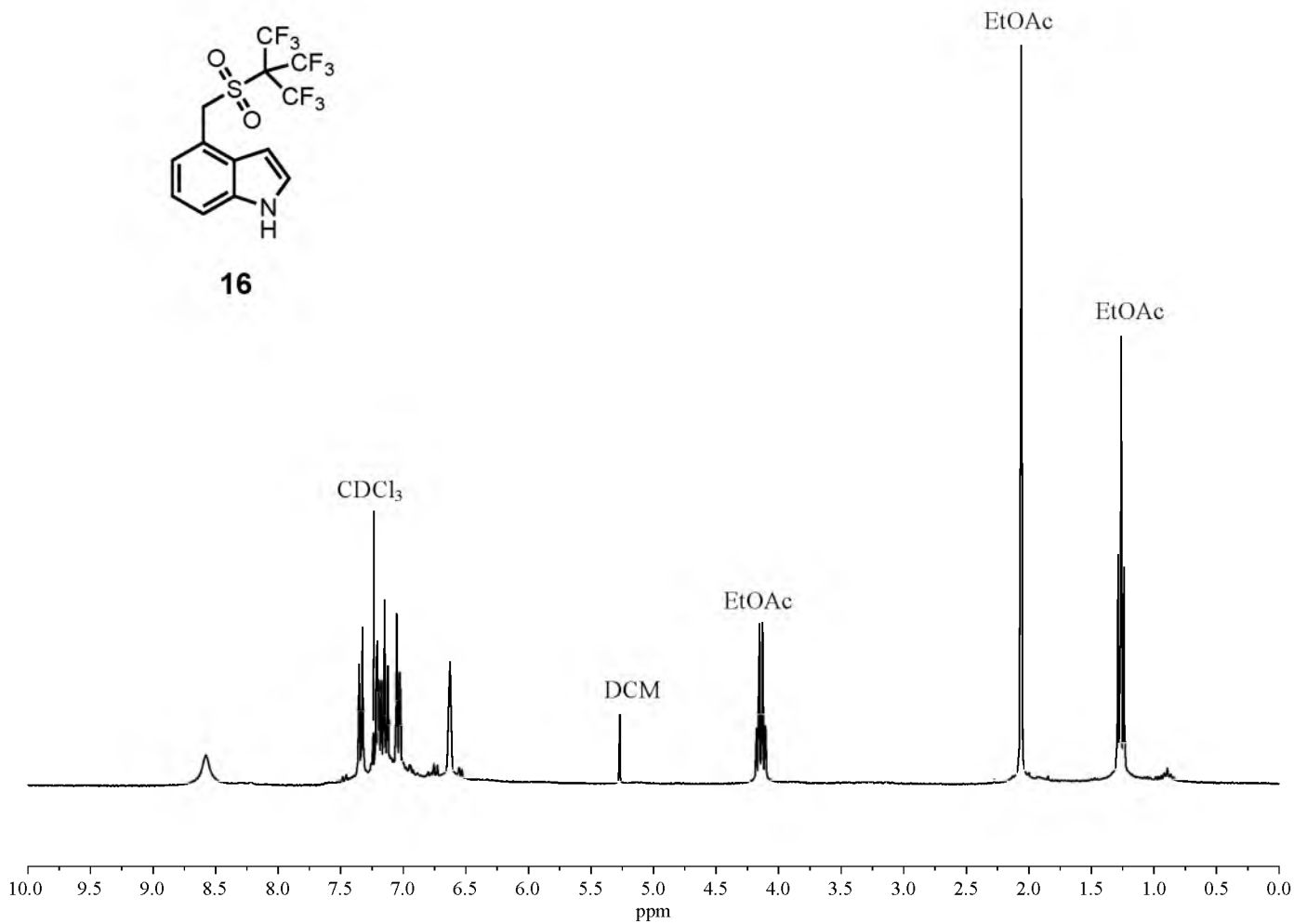


NMR 11. 300 MHz ^{1}H NMR spectrum of compound 15 in D_2O with DSS.

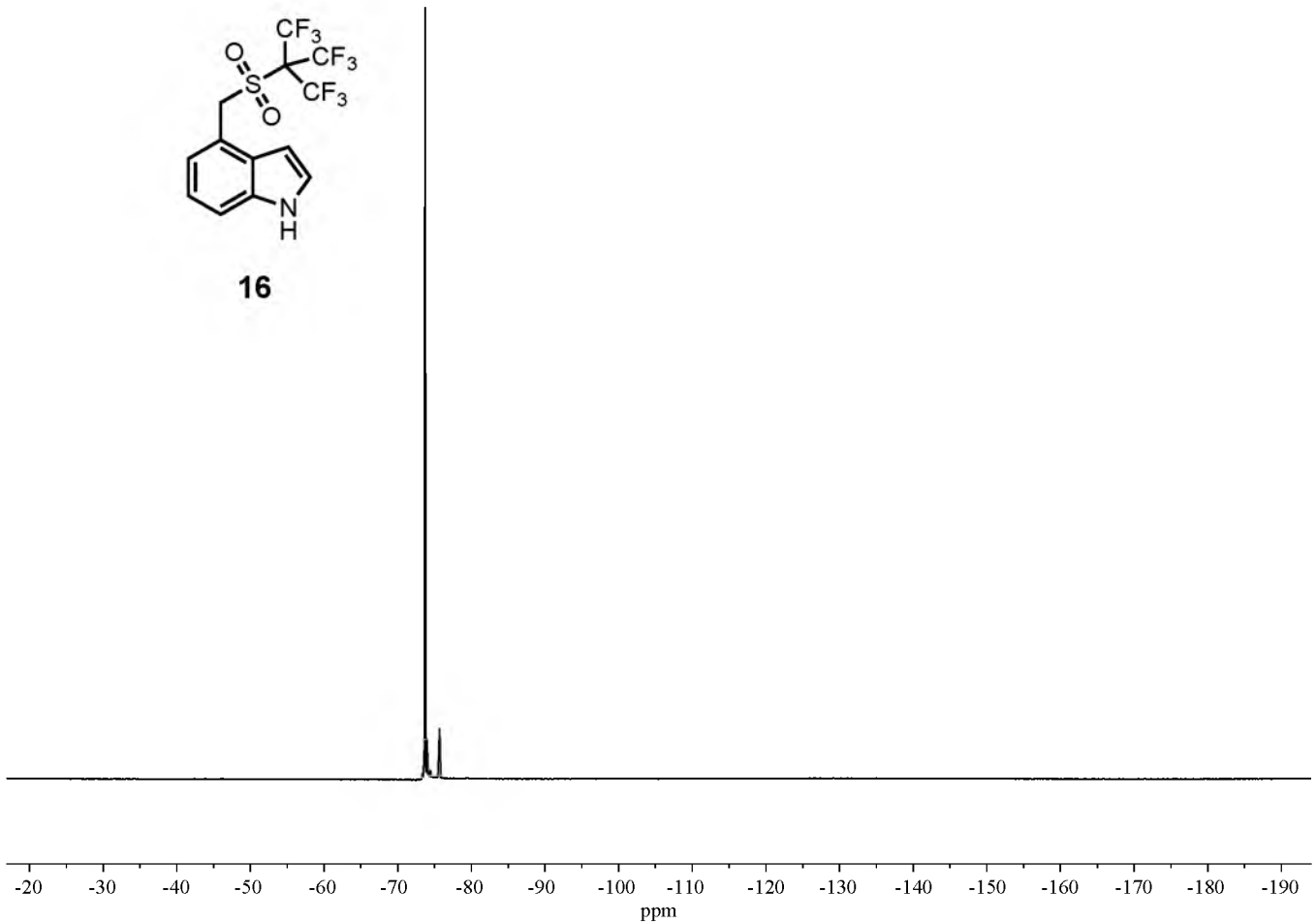


NMR 12. 75 MHz ^{13}C NMR spectrum of compound 15 in D_2O.

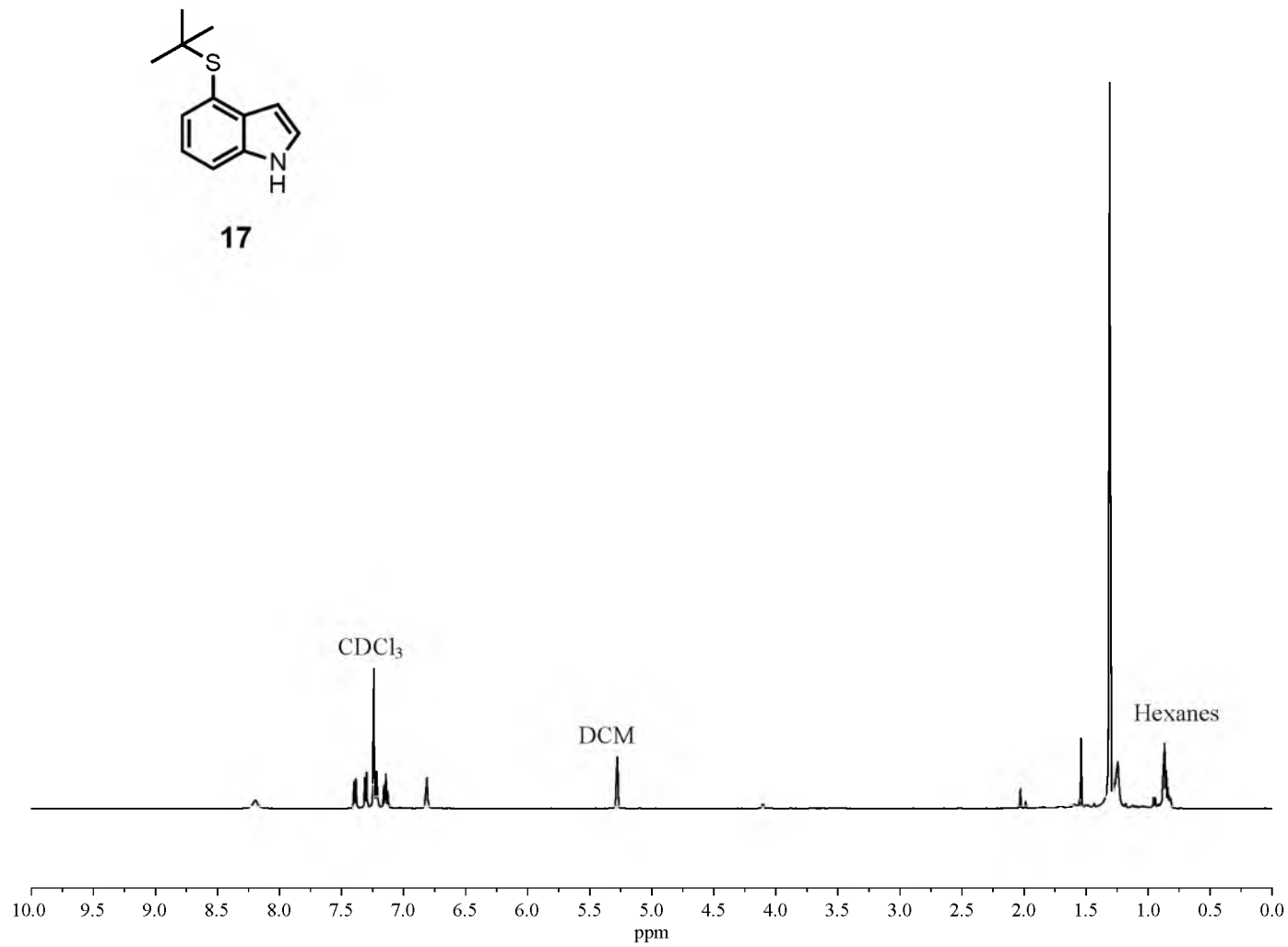

0
10
20
30
40
50
60
70
80



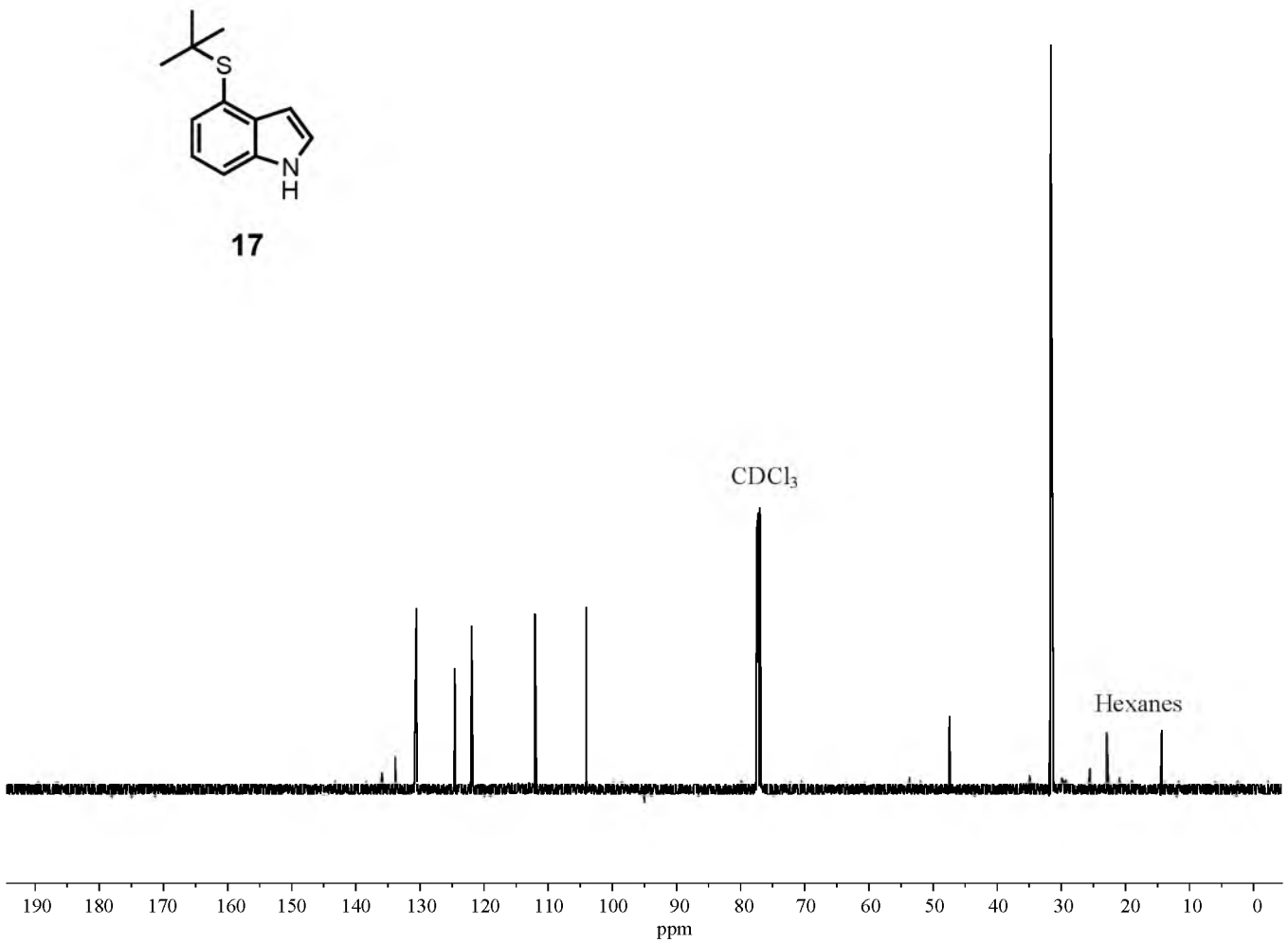
NMR 13. 500 MHz ^{1}H NMR spectrum of compound 16 in $CDCl_3$.



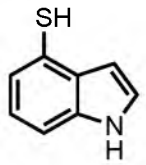
NMR 14. 470 MHz ^{19}F NMR spectrum of compound 16 in $CDCl_3$.



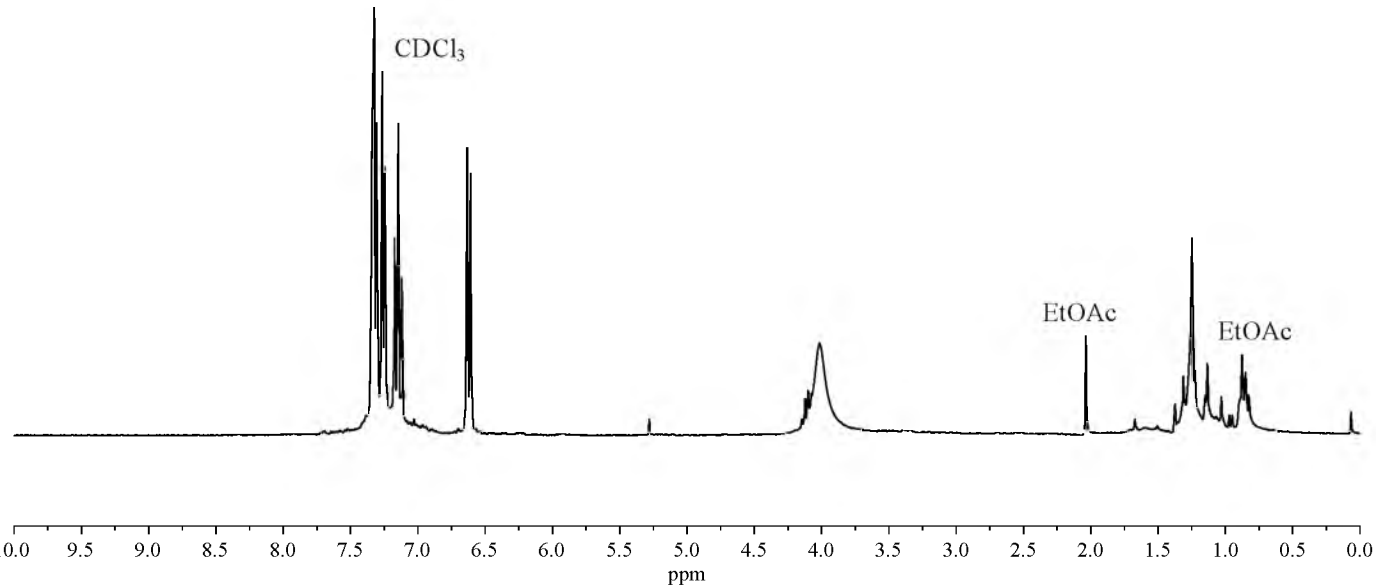
NMR 15. 500 MHz ^{1}H NMR spectrum of compound 17 in $CDCl_3$.



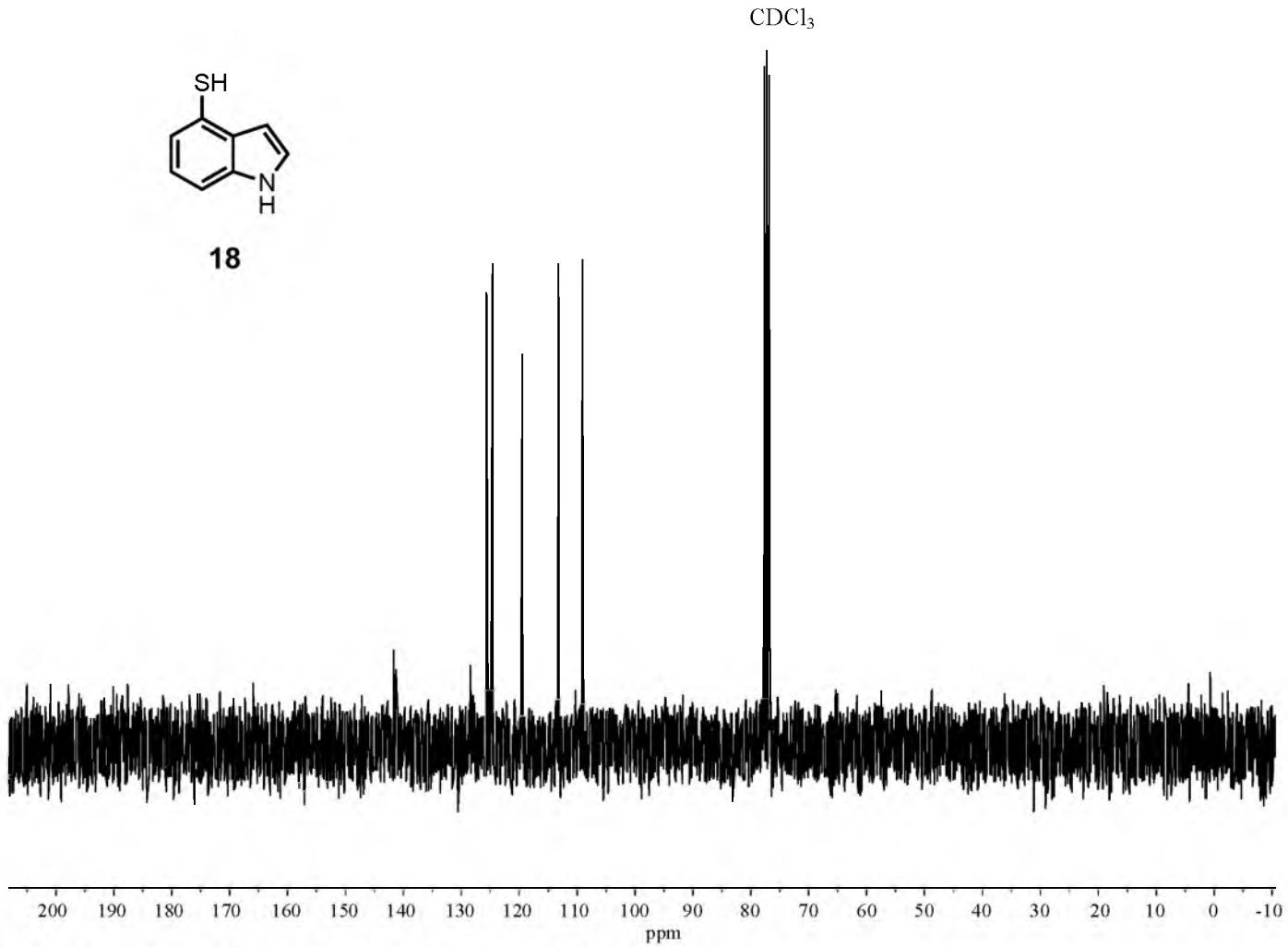
NMR 16. 125 MHz ^{13}C NMR spectrum of compound 17 in $CDCl_3$.



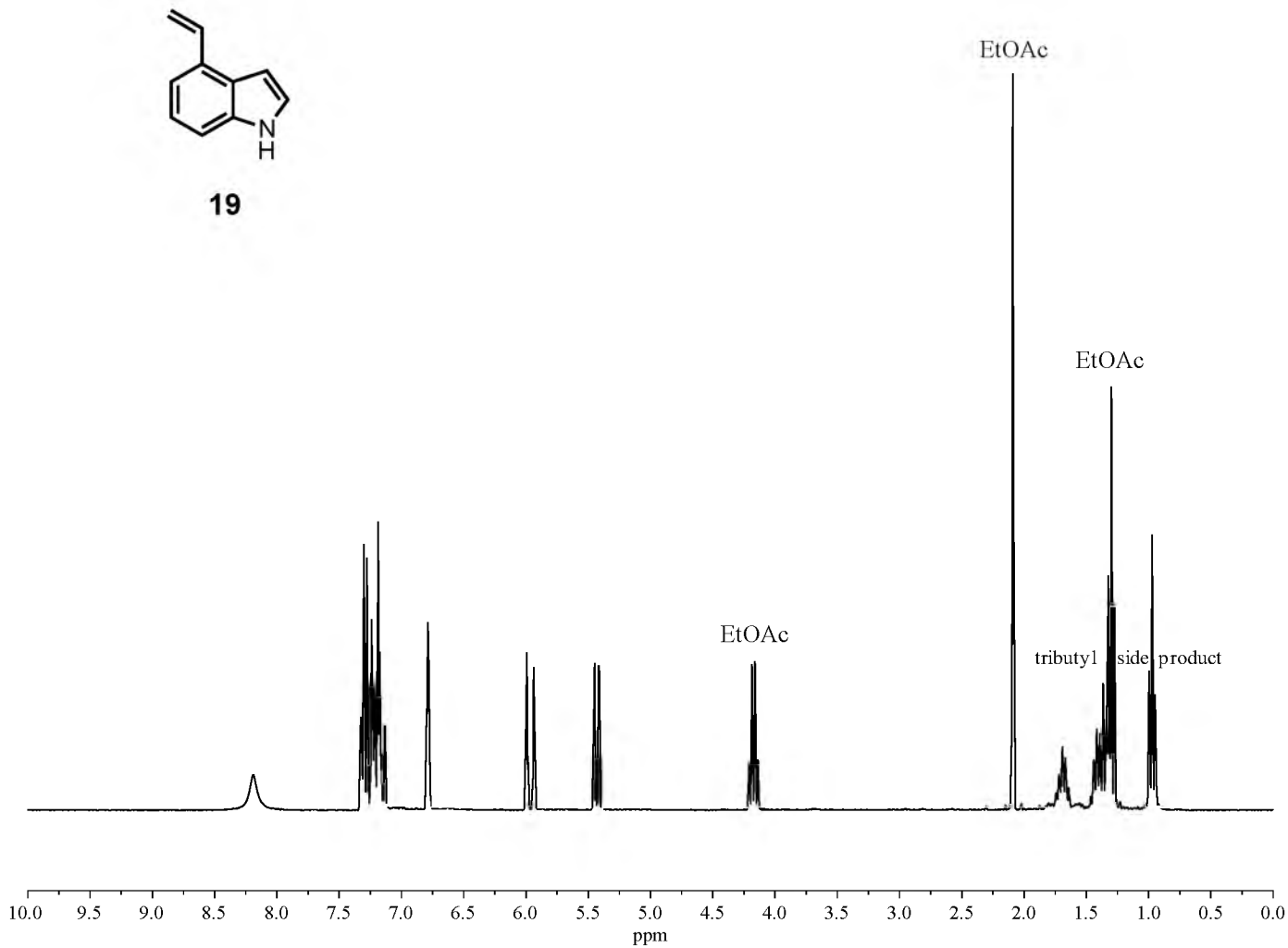
18



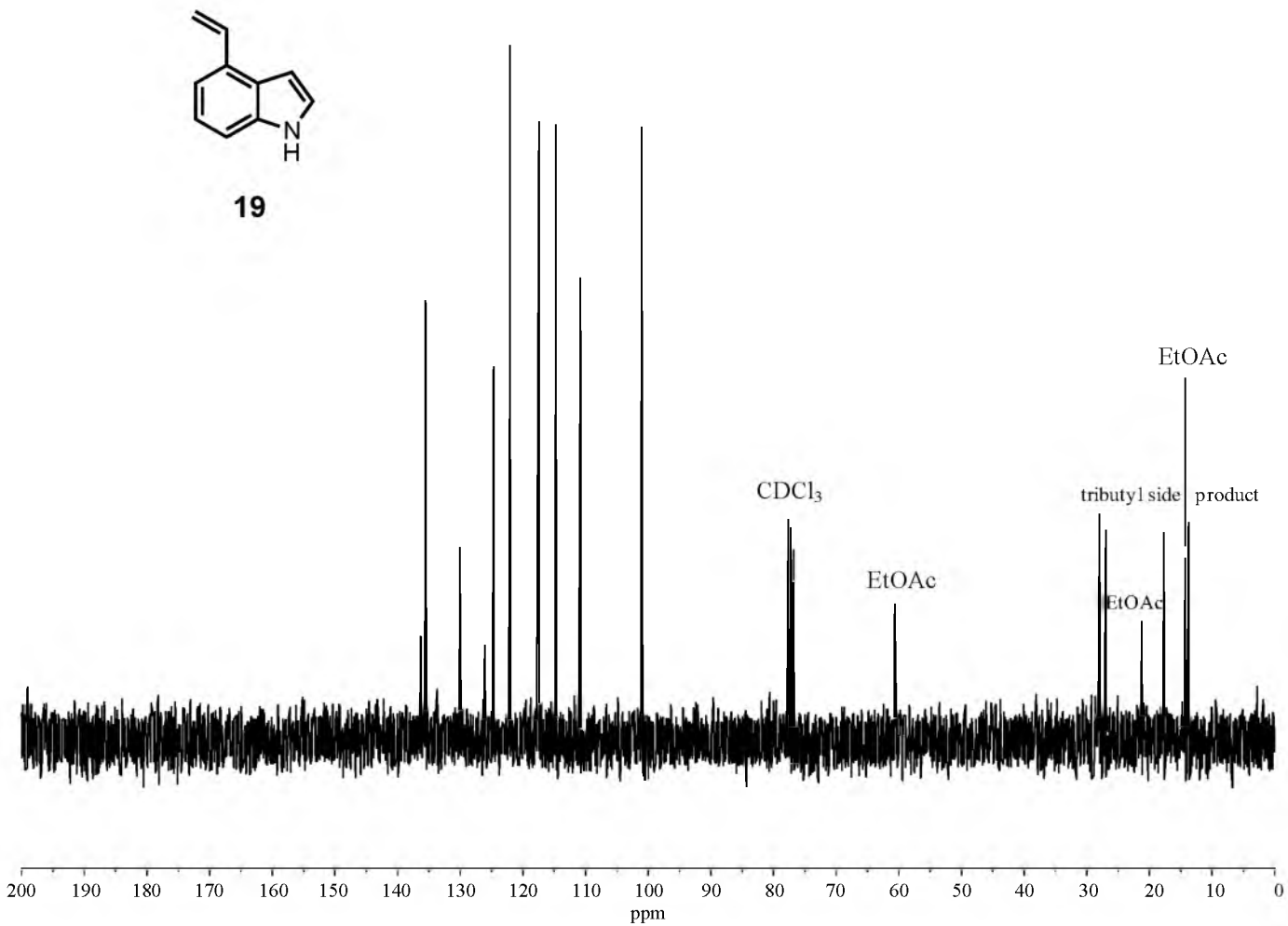
NMR 17. 300 MHz ^{1}H NMR spectrum of compound 18 in $CDCl_3$.



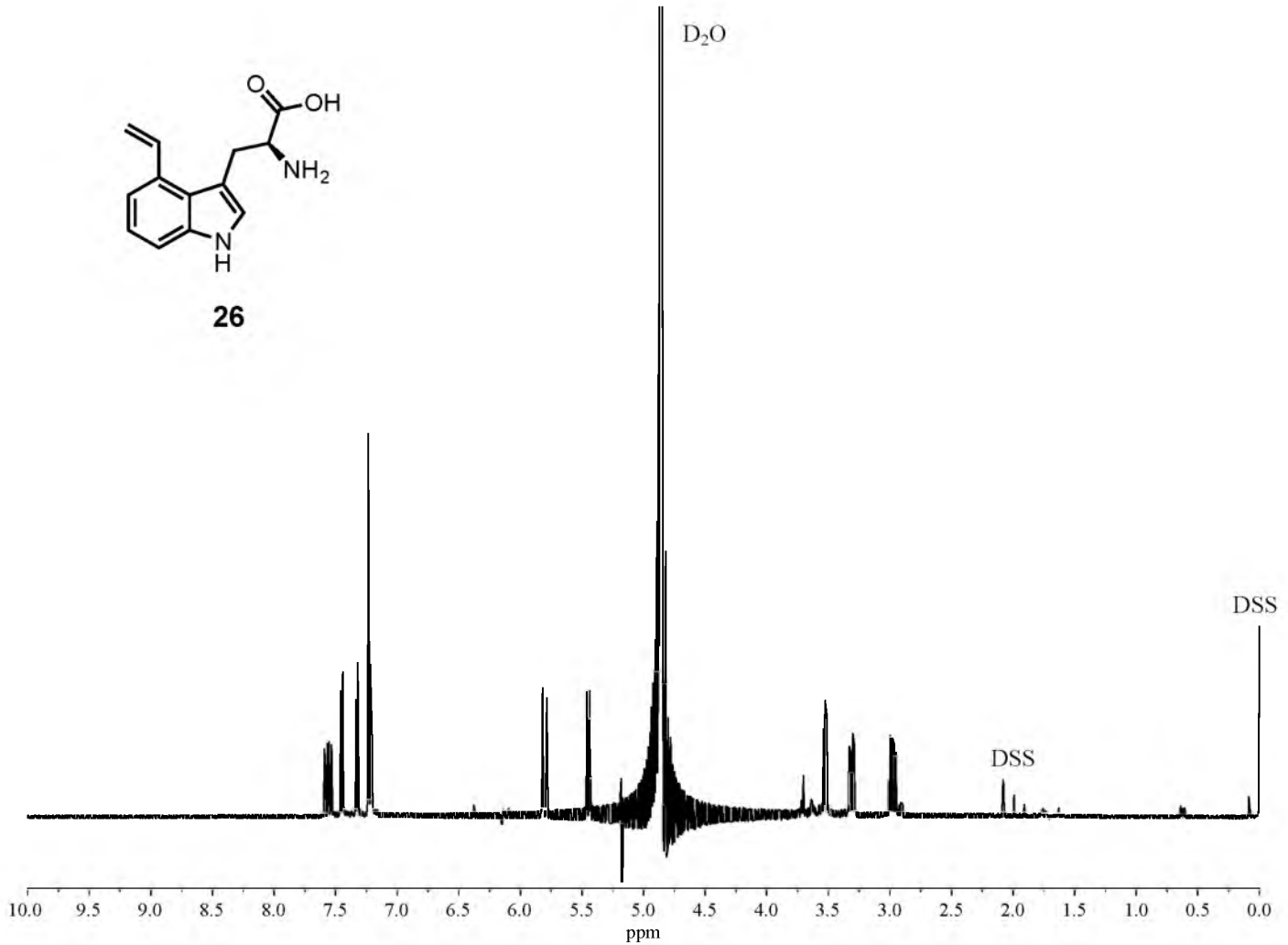
NMR 18. 75 MHz ^{13}C NMR spectrum of compound 18 in $CDCl_3$.



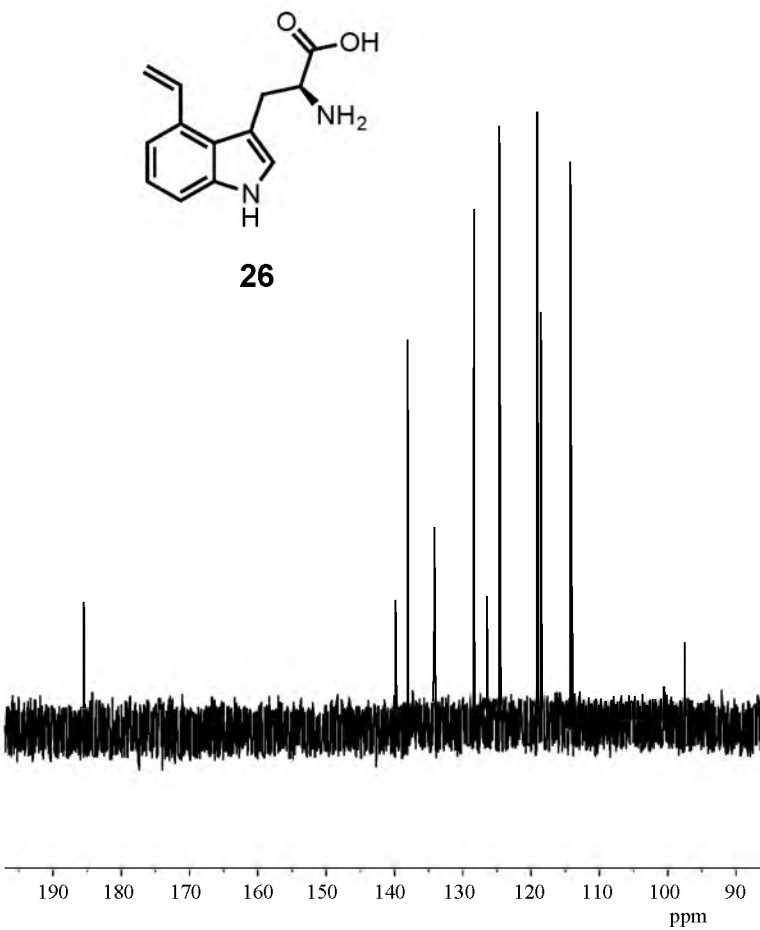
NMR 19. 500 MHz ^{1}H NMR spectrum of compound 19 in $CDCl_3$.



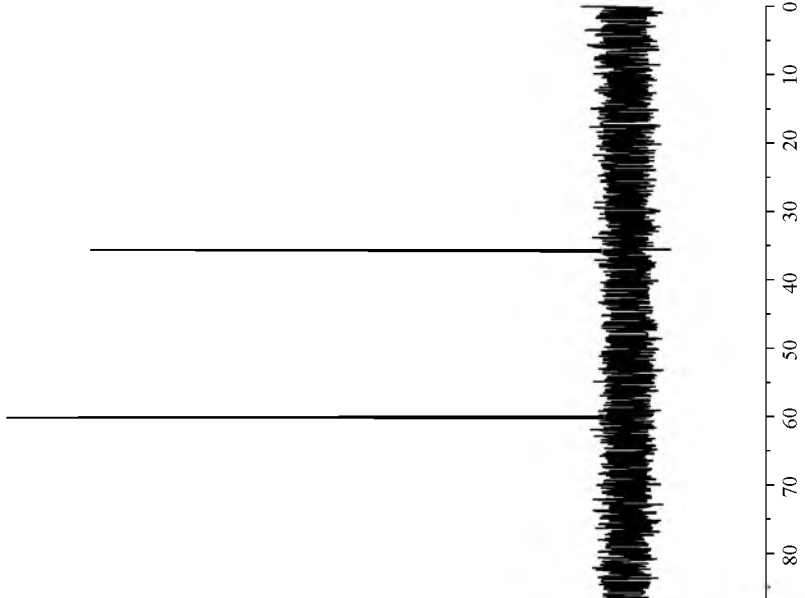
NMR 20. 125 MHz ^{13}C NMR spectrum of compound 19 in $CDCl_3$.

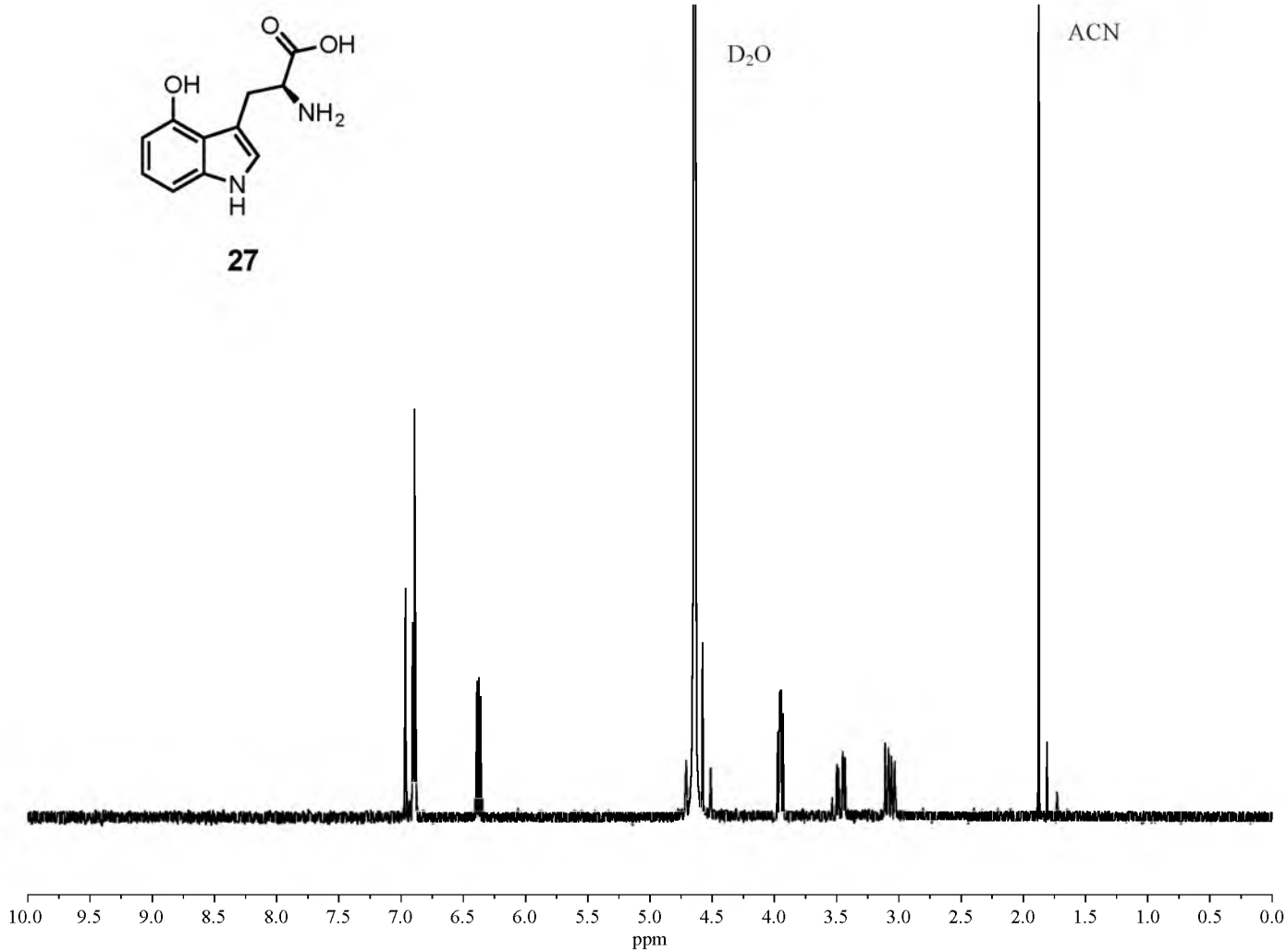


NMR 21. 500 MHz ^{1}H NMR spectrum of compound 26 in D_2O with DSS.

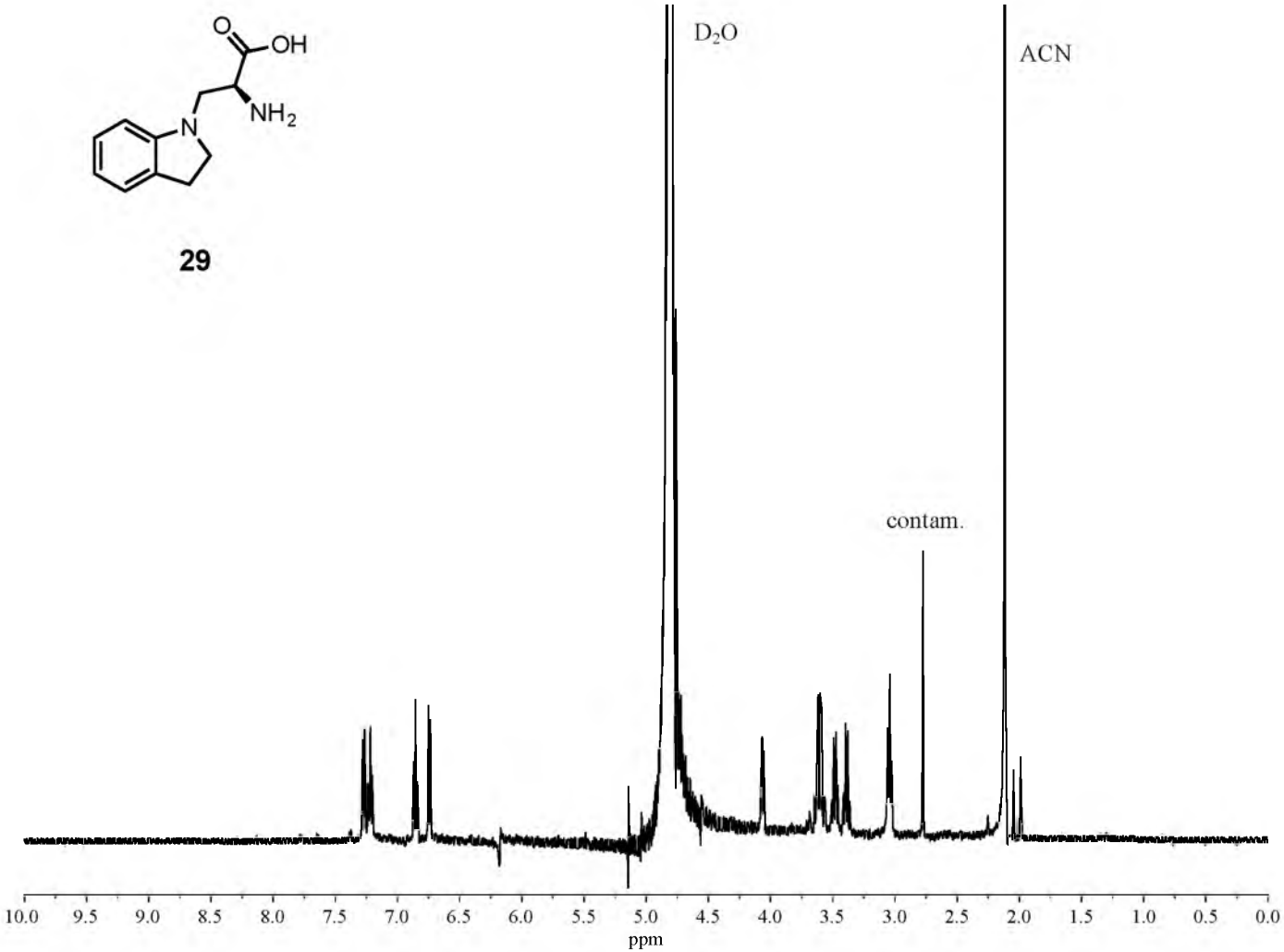


NMR 22. 125 MHz ^{13}C NMR spectrum of compound 26 in D_2O.

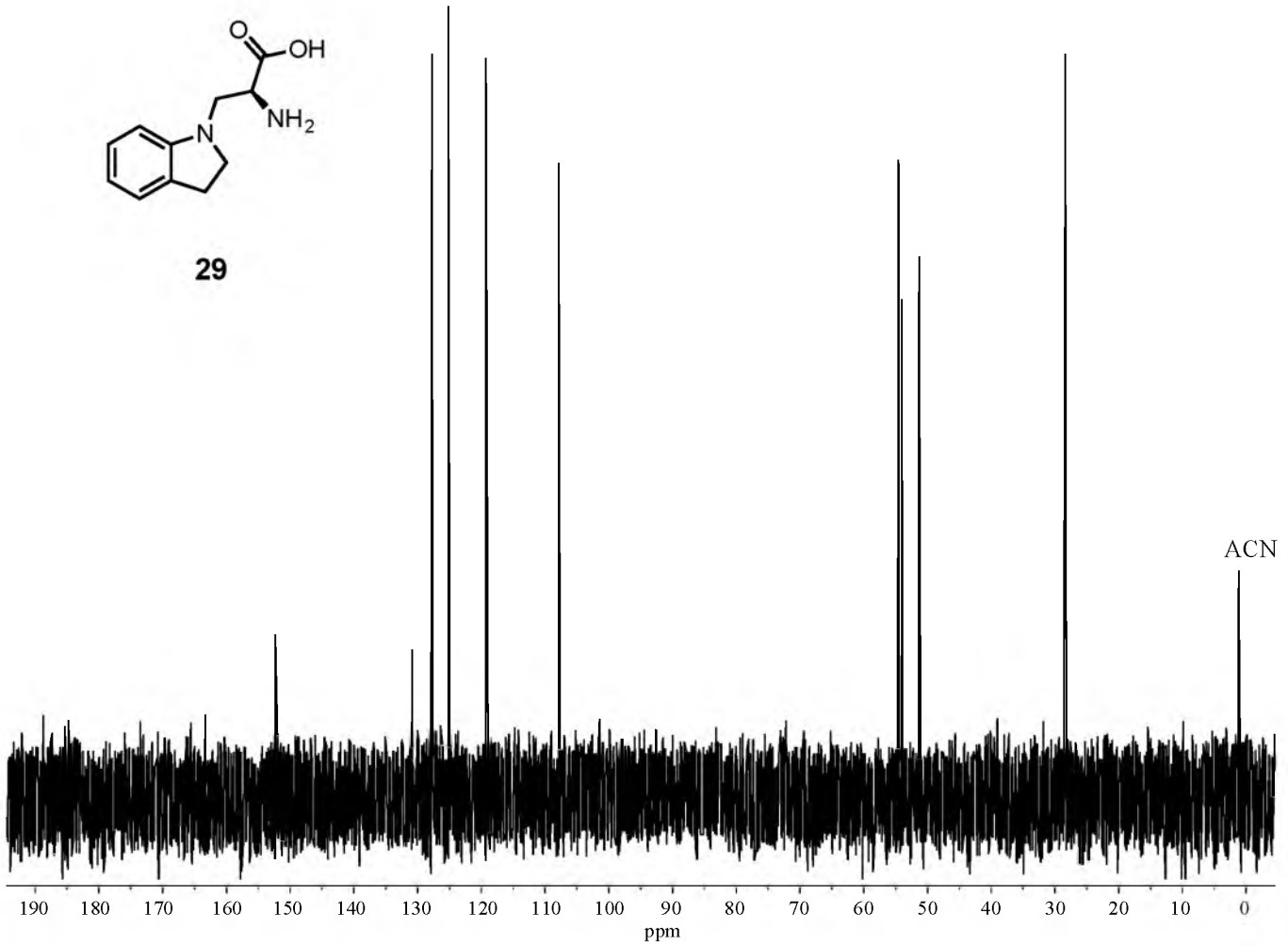

0
10
20
30
40
50
60
70
80



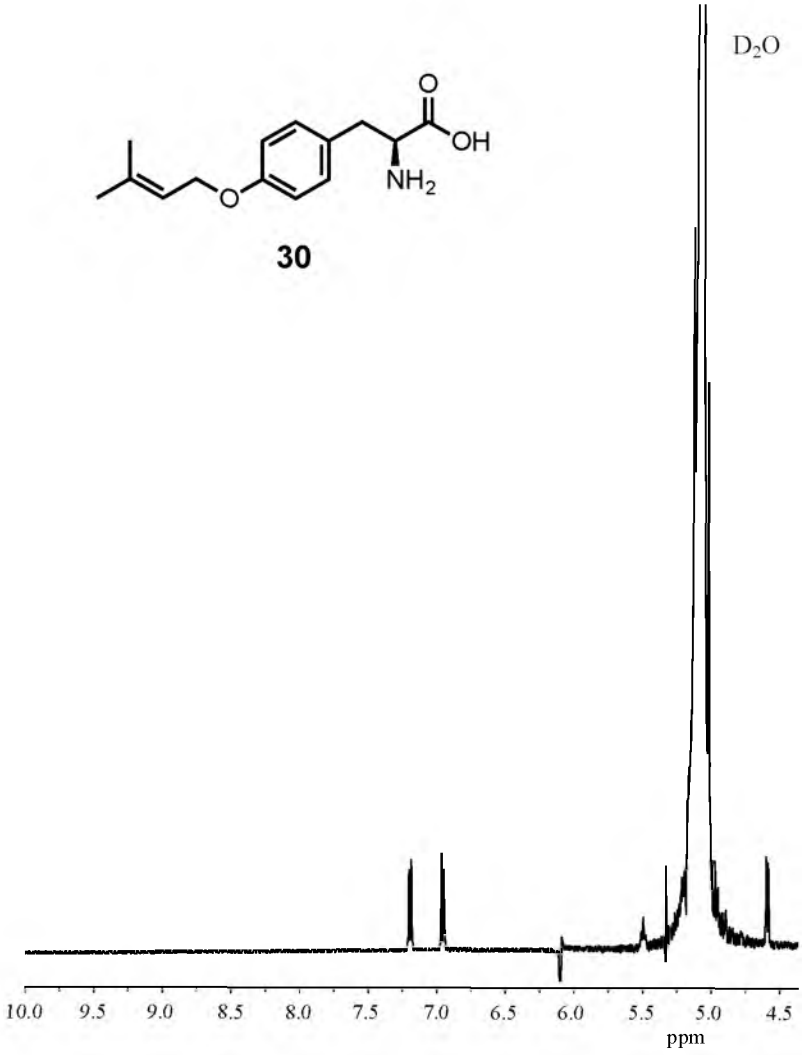
NMR 23. 300 MHz ^{1}H NMR spectrum of compound 27 in D_2O.



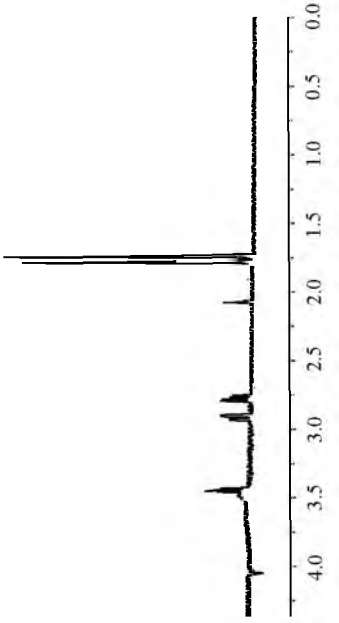
NMR 24. 500 MHz ^{1}H NMR spectrum of compound 29 in D_2O.

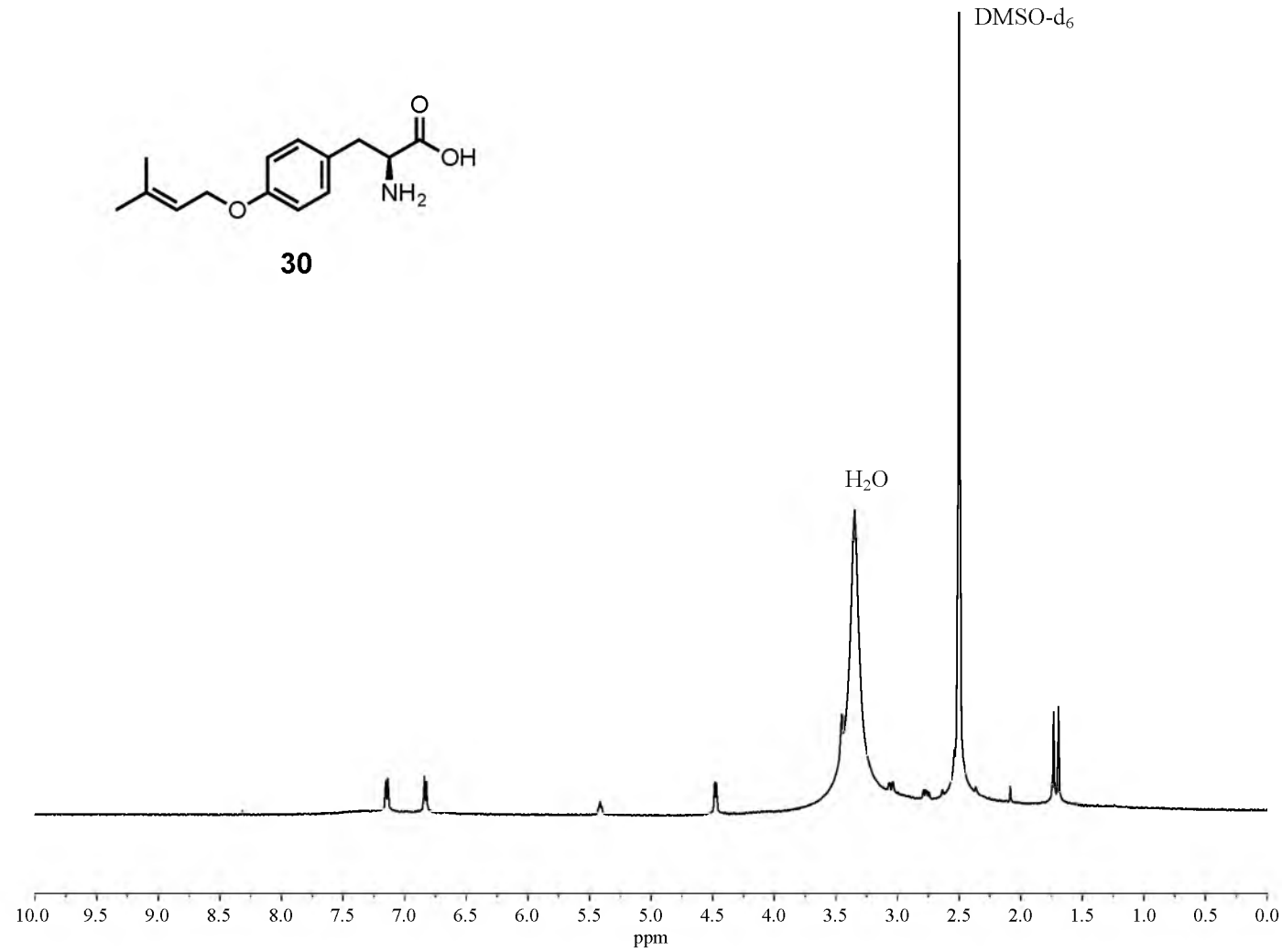


NMR 25. 125 MHz ^{13}C NMR spectrum of compound 29 in D_2O.

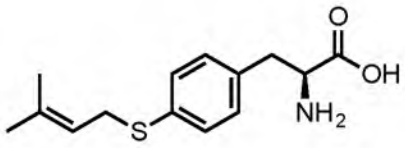


NMR 26. 500 MHz ^{1}H NMR spectrum of compound 30 in D_2O.

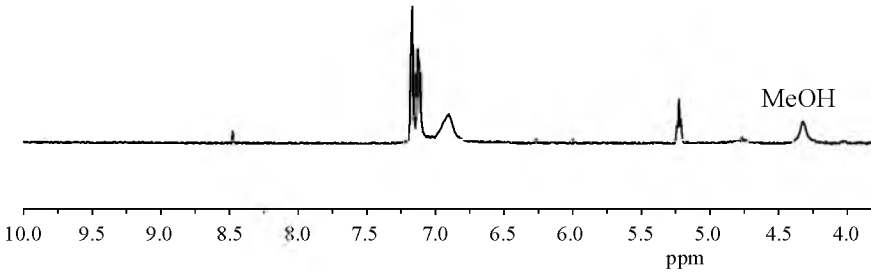

4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0



NMR 27. 500 MHz ^{1}H NMR spectrum of compound 30 in DMSO-d$_6$.

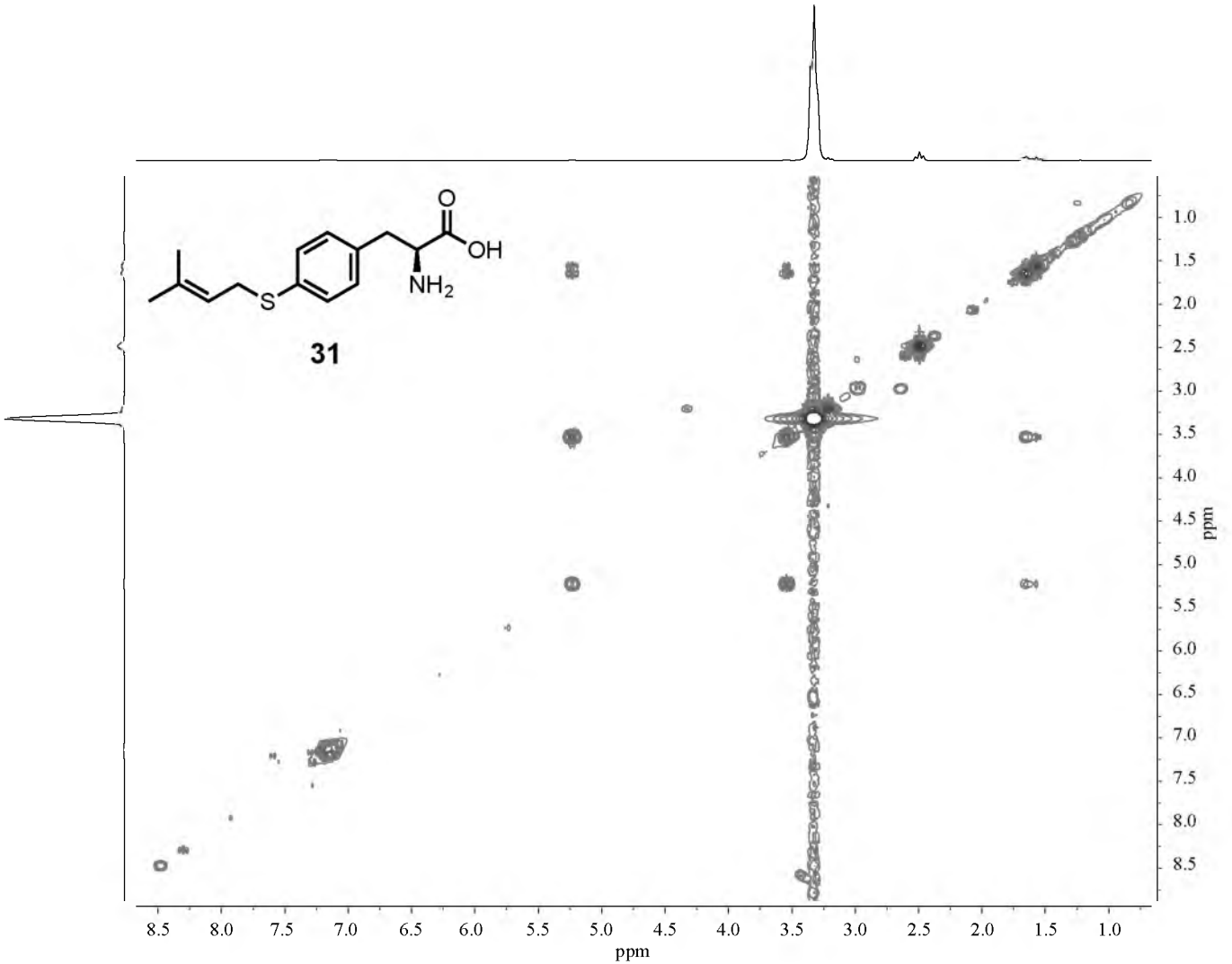


31

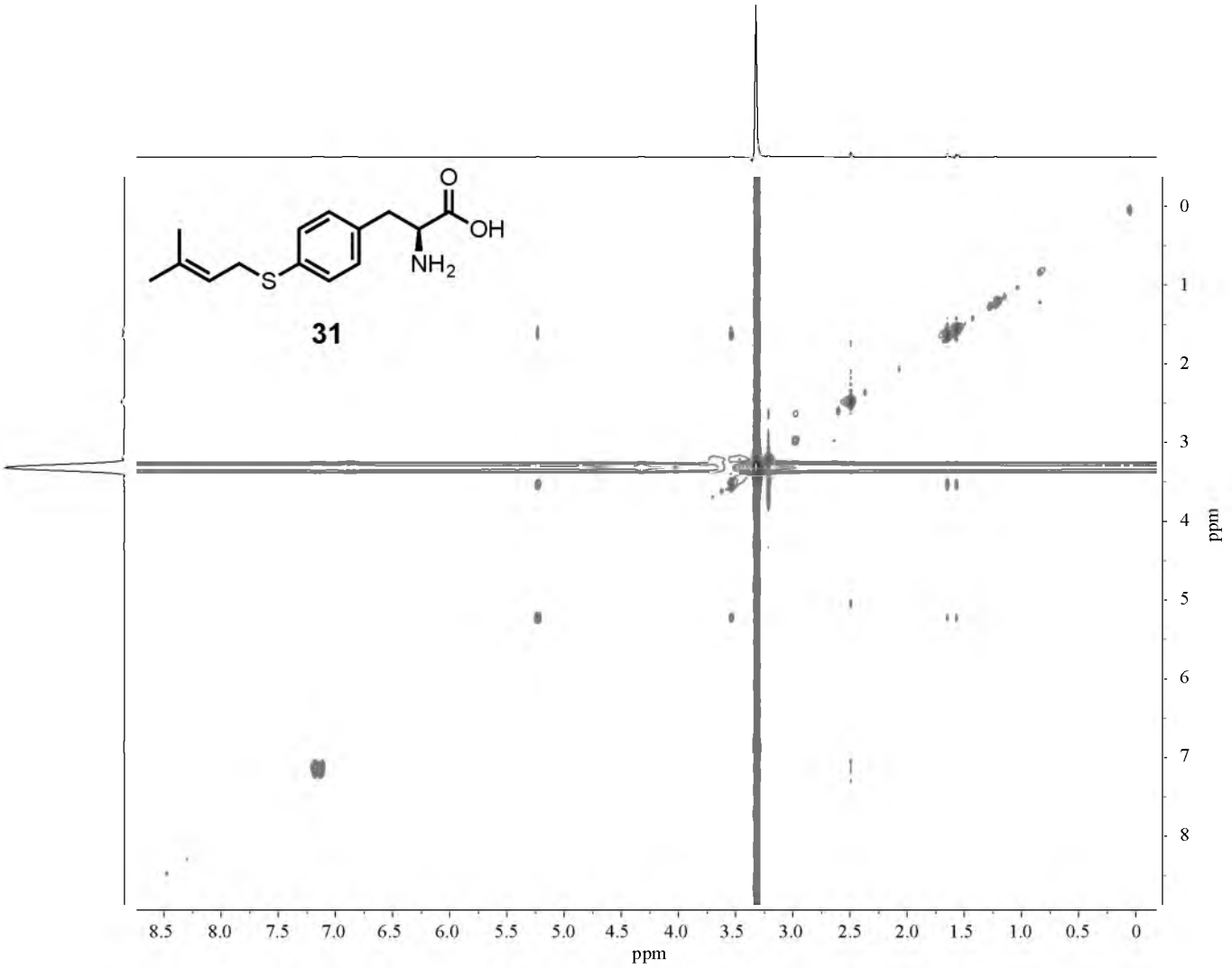


NMR 28. 600 MHz ^{1}H NMR spectrum of compound 31 in DMSO-d$_6$.

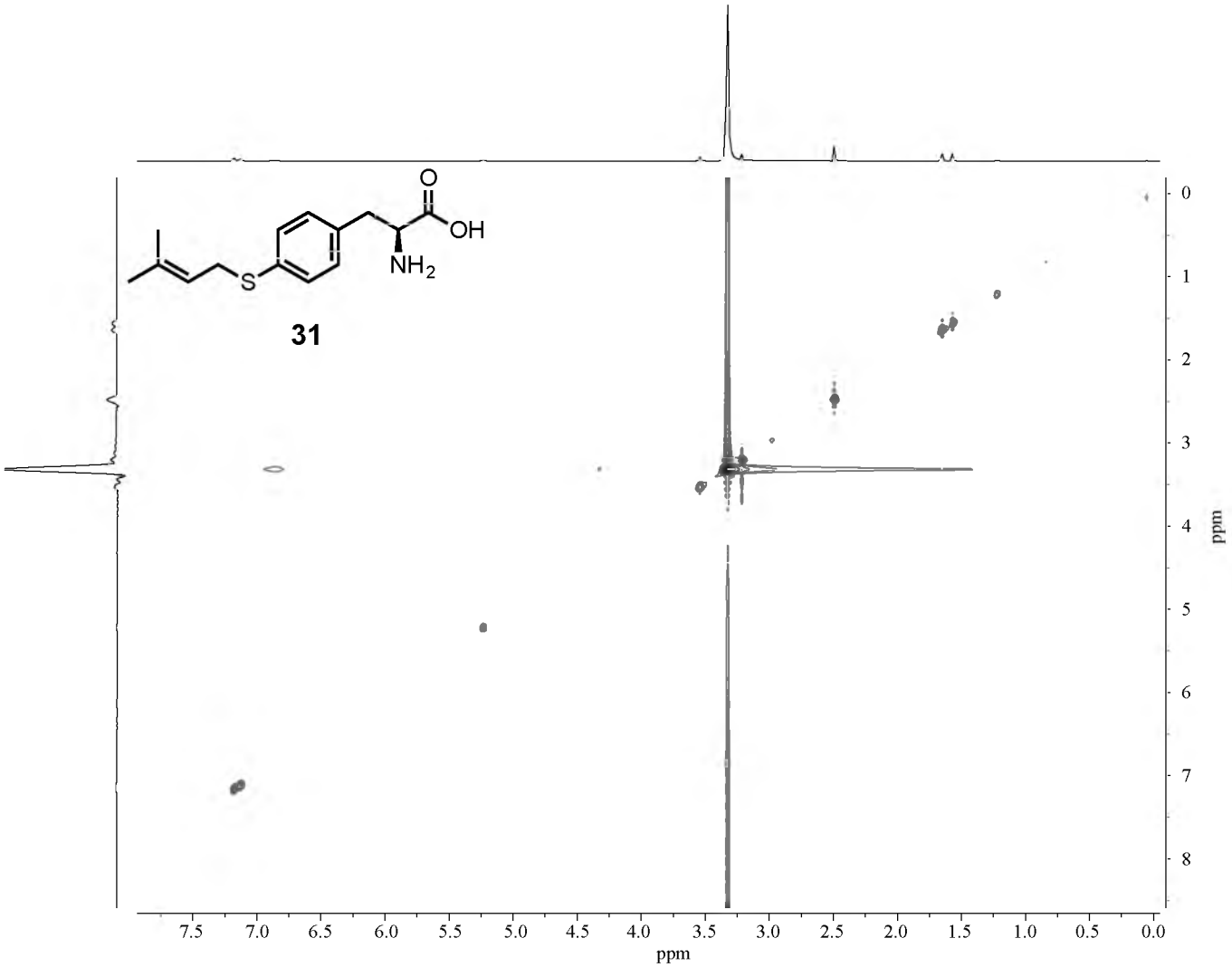
H2O
DMSO-d6
DMSO-d6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0



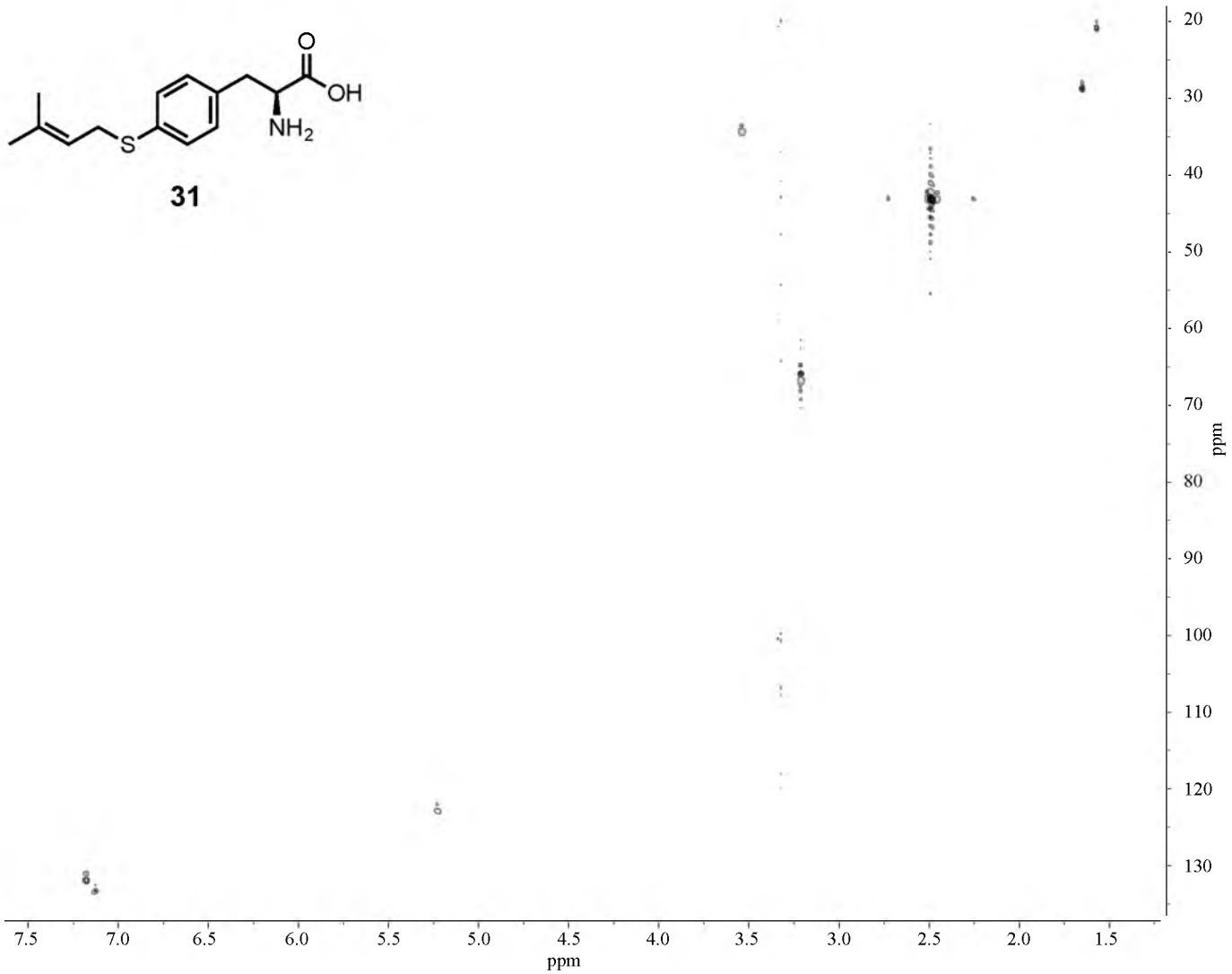
NMR 29. ^{1}H-^{1}H COSY NMR spectrum of compound 31 in DMSO-d_6.



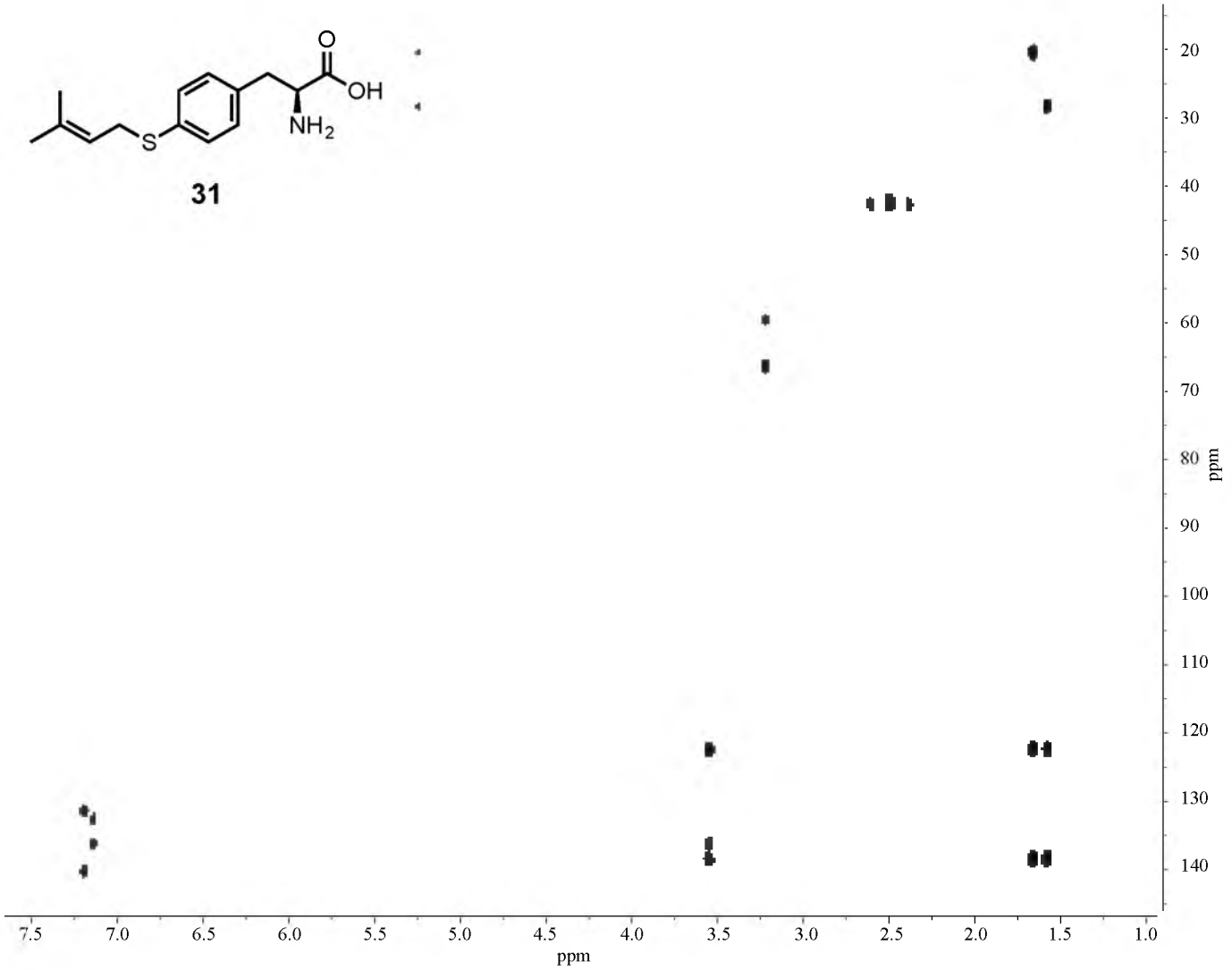
NMR 30. ^{1}H-^{1}H TOCSY NMR spectrum of compound 31 in DMSO-d_6.



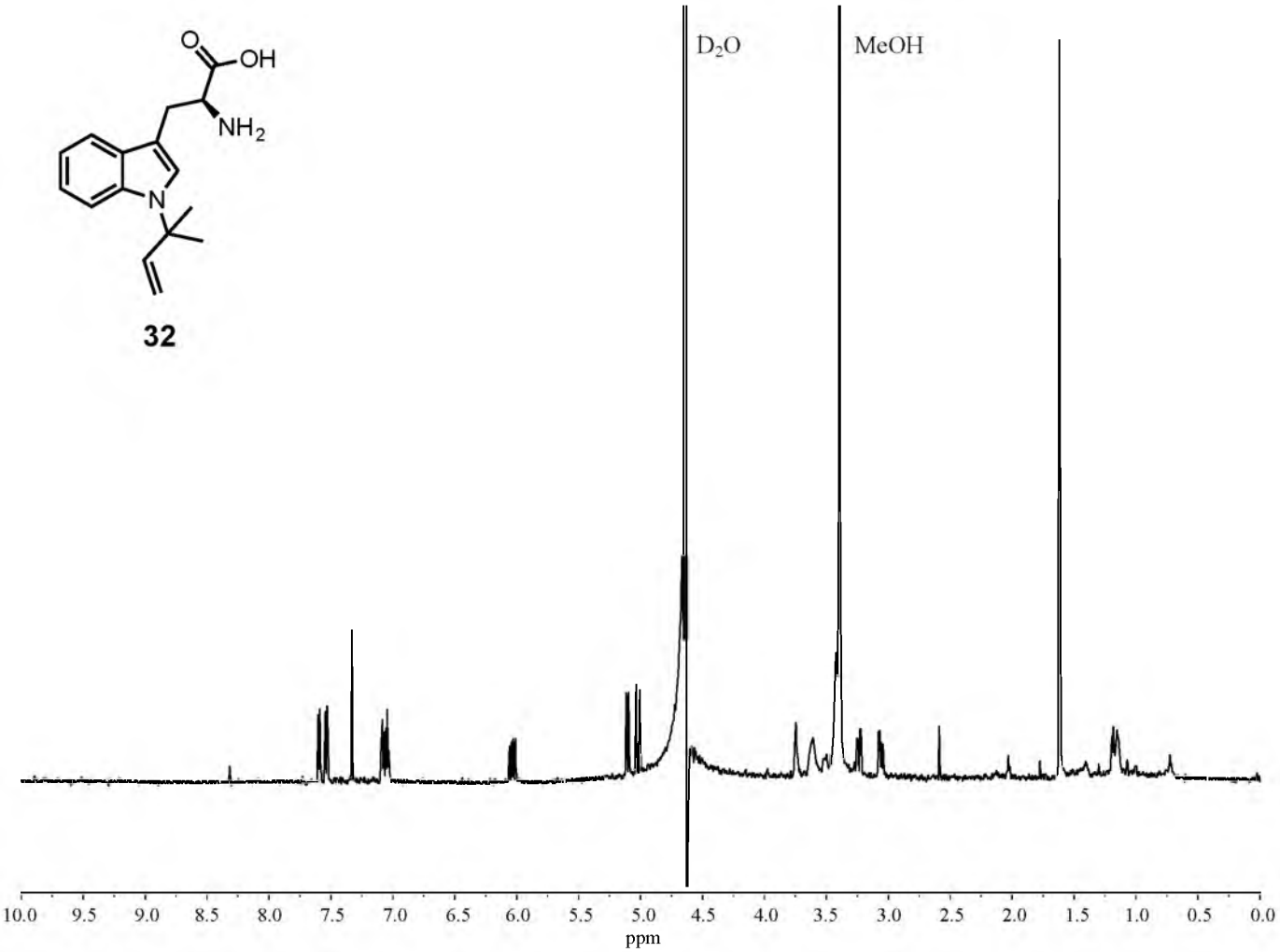
NMR 31. ^{1}H-^{1}H ROESY NMR spectrum of compound 31 in DMSO-d_6.



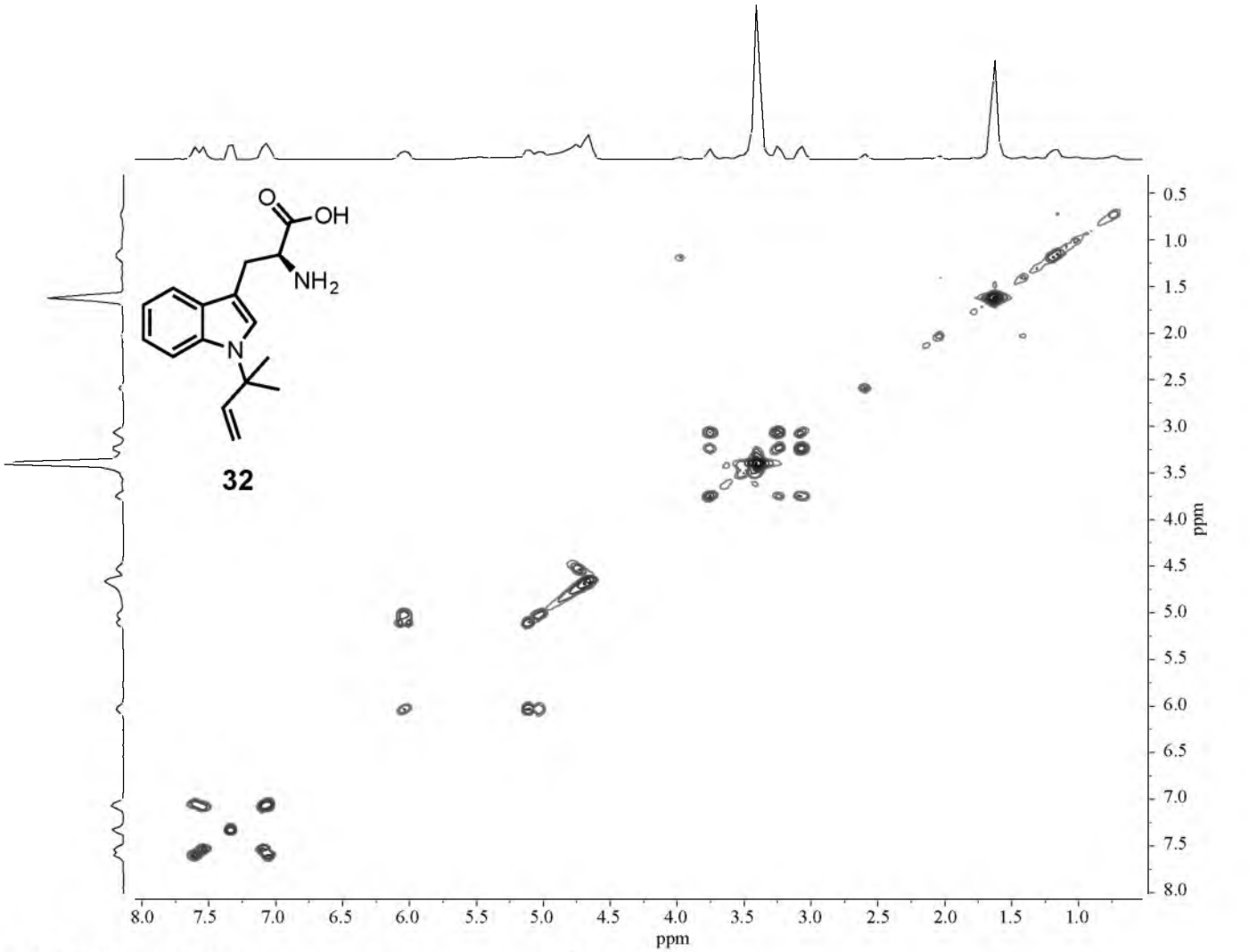
NMR 32. ^{1}H-^{13}C HSQC NMR spectrum of compound 31 in DMSO-d_6.



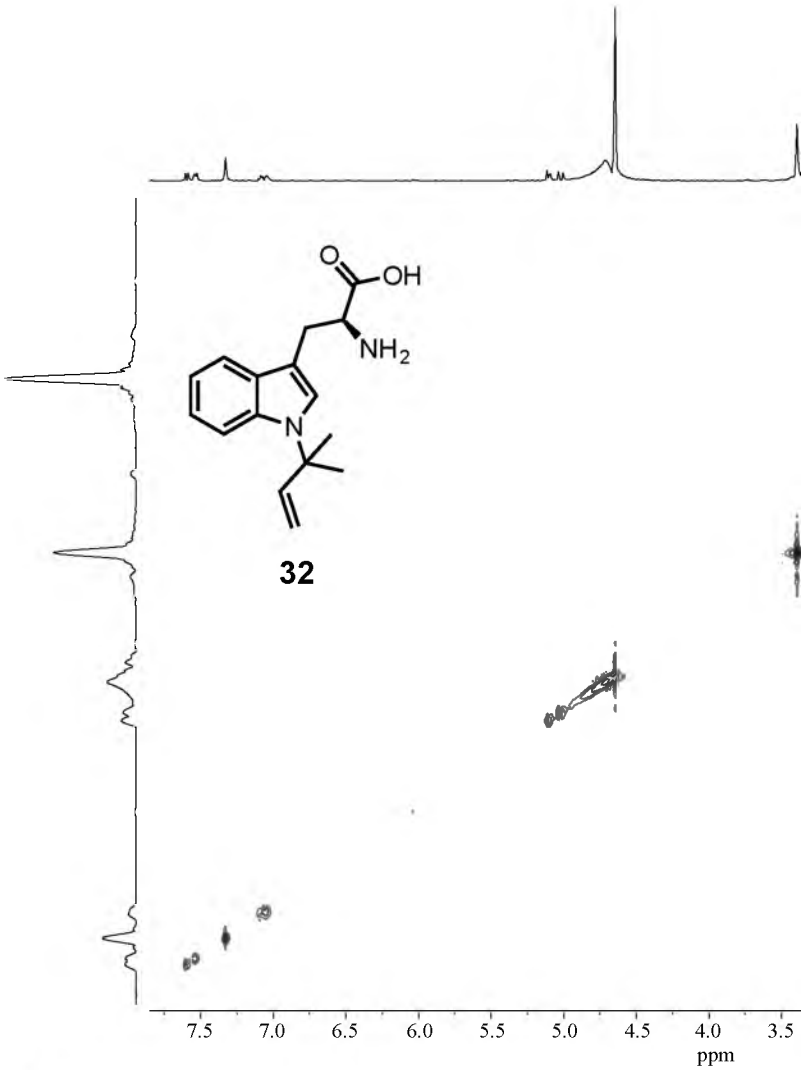
NMR 33. ^{1}H-^{13}C HMBC NMR spectrum of compound 31 in DMSO-d_6.



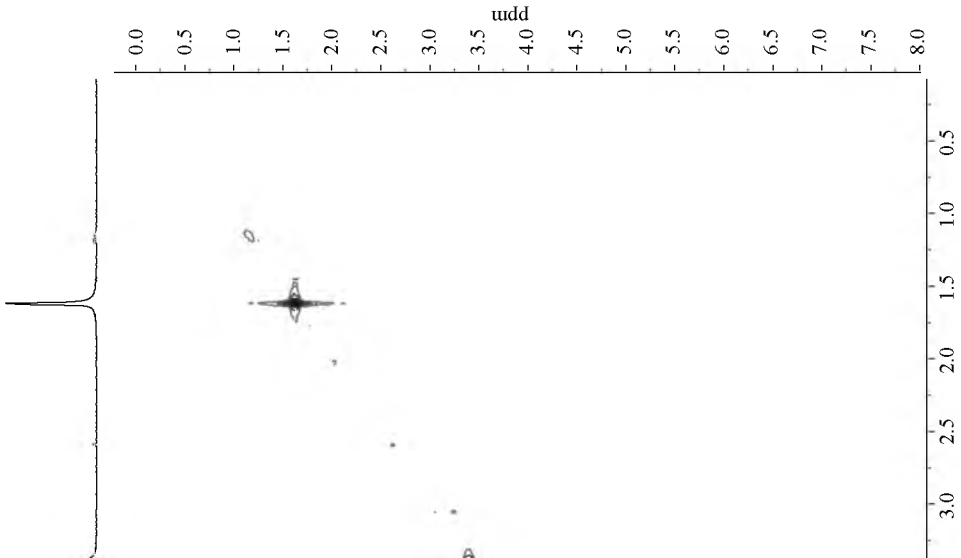
NMR 34. 600 MHz ^{1}H NMR spectrum of compound 32 in D_2O.

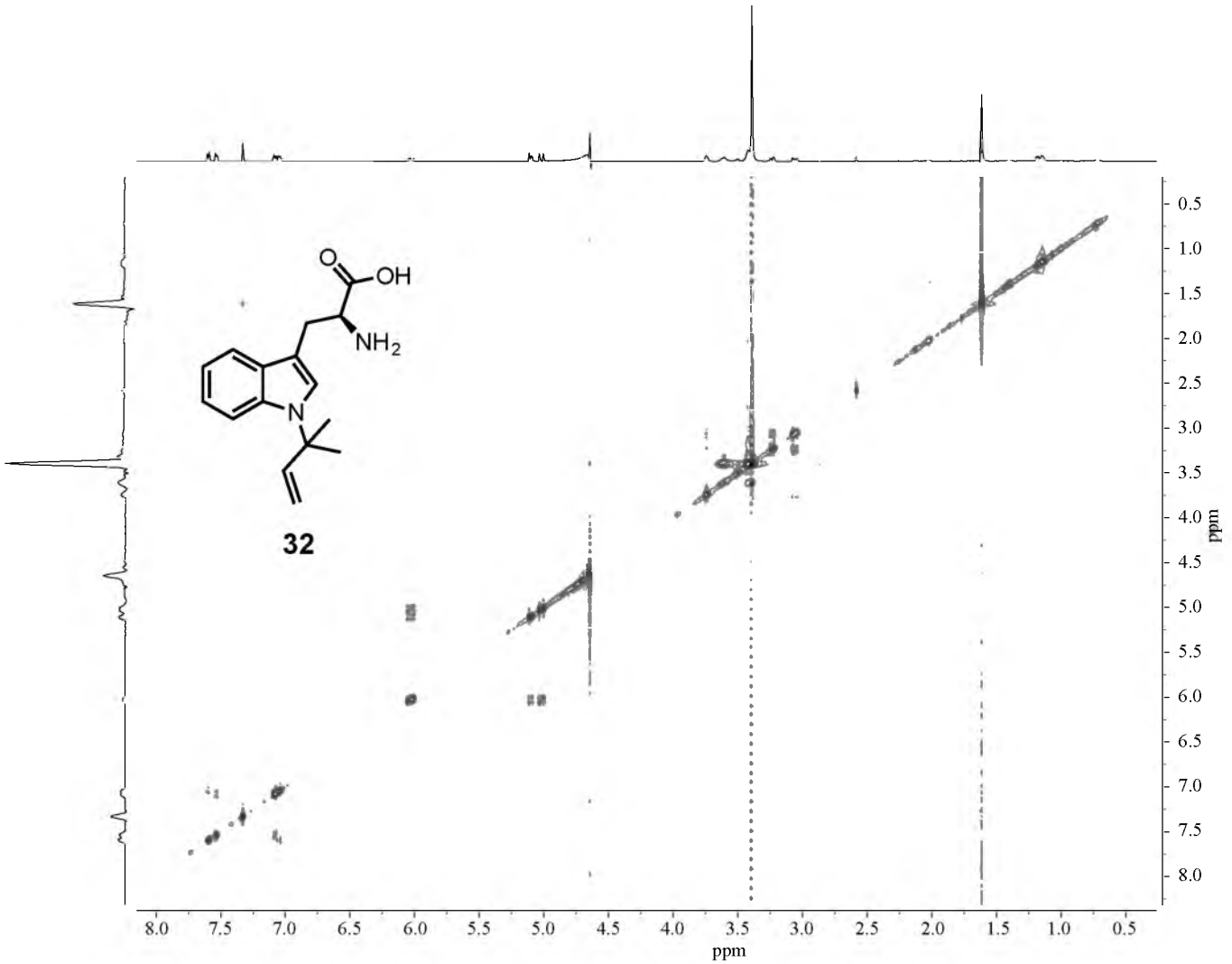


NMR 35. ^{1}H-^{1}H COSY NMR spectrum of compound 32 in D_2O.

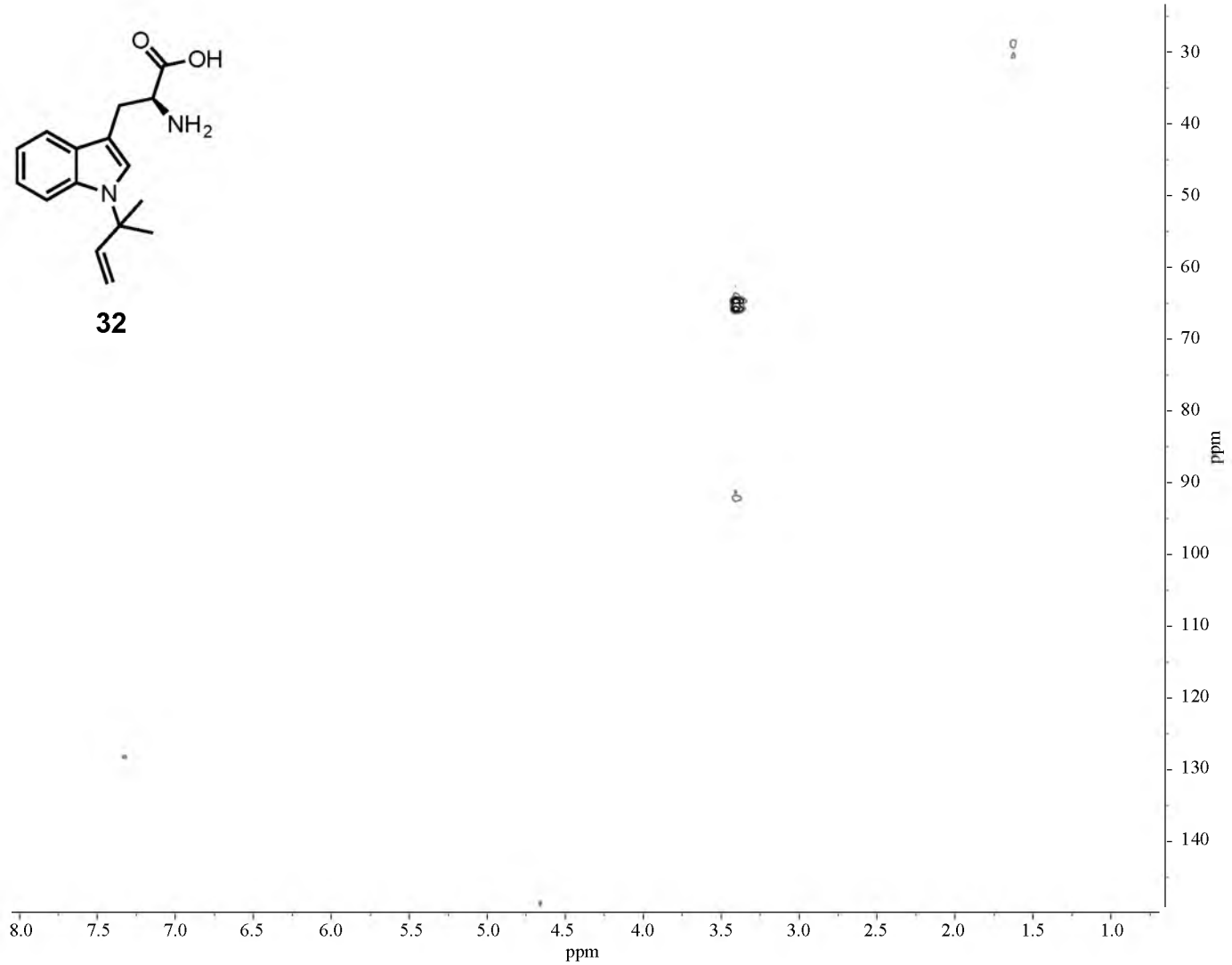


NMR 36. ^{1}H-^{1}H TOCSY NMR spectrum of compound 32 in D_2O.

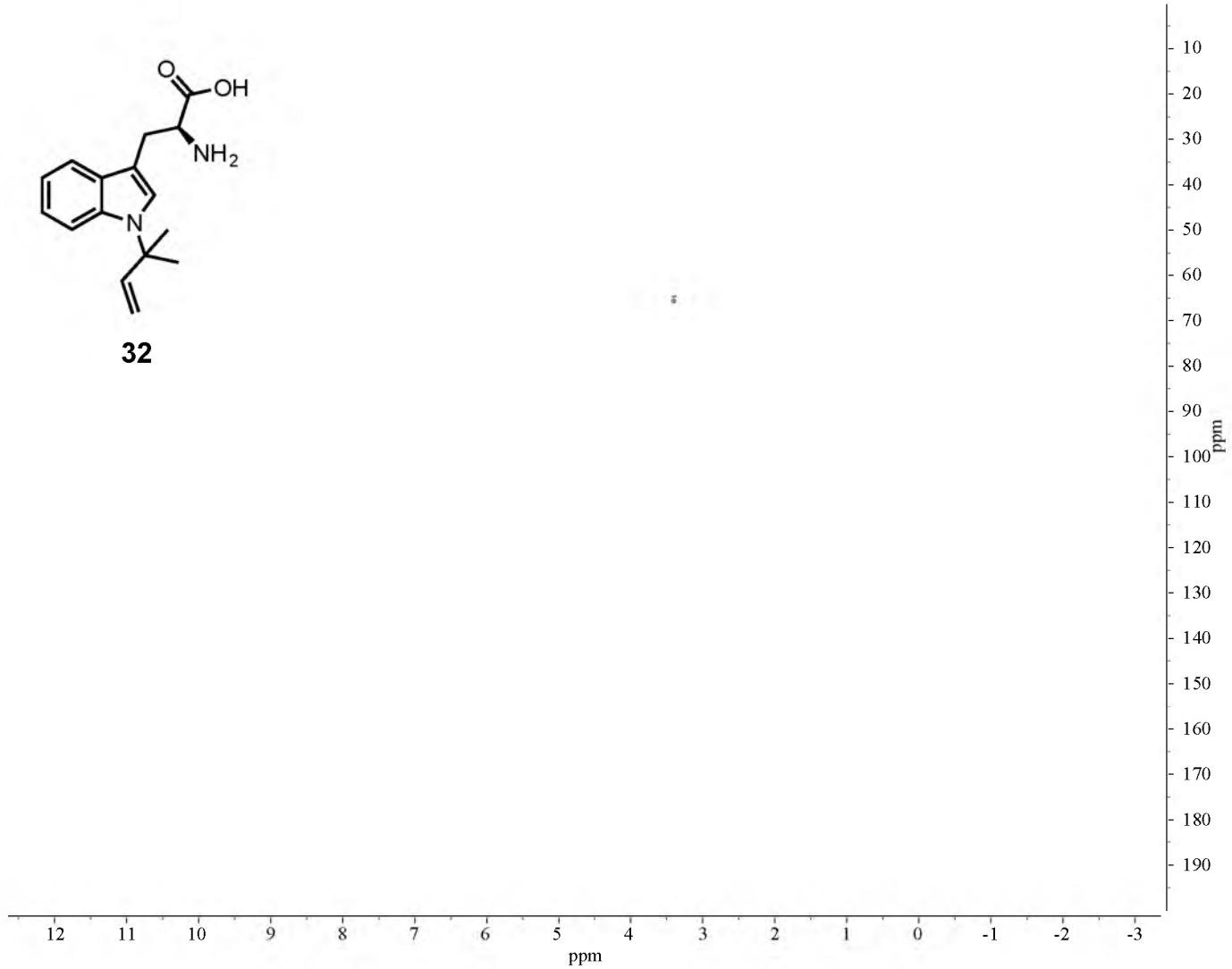

ppm
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0



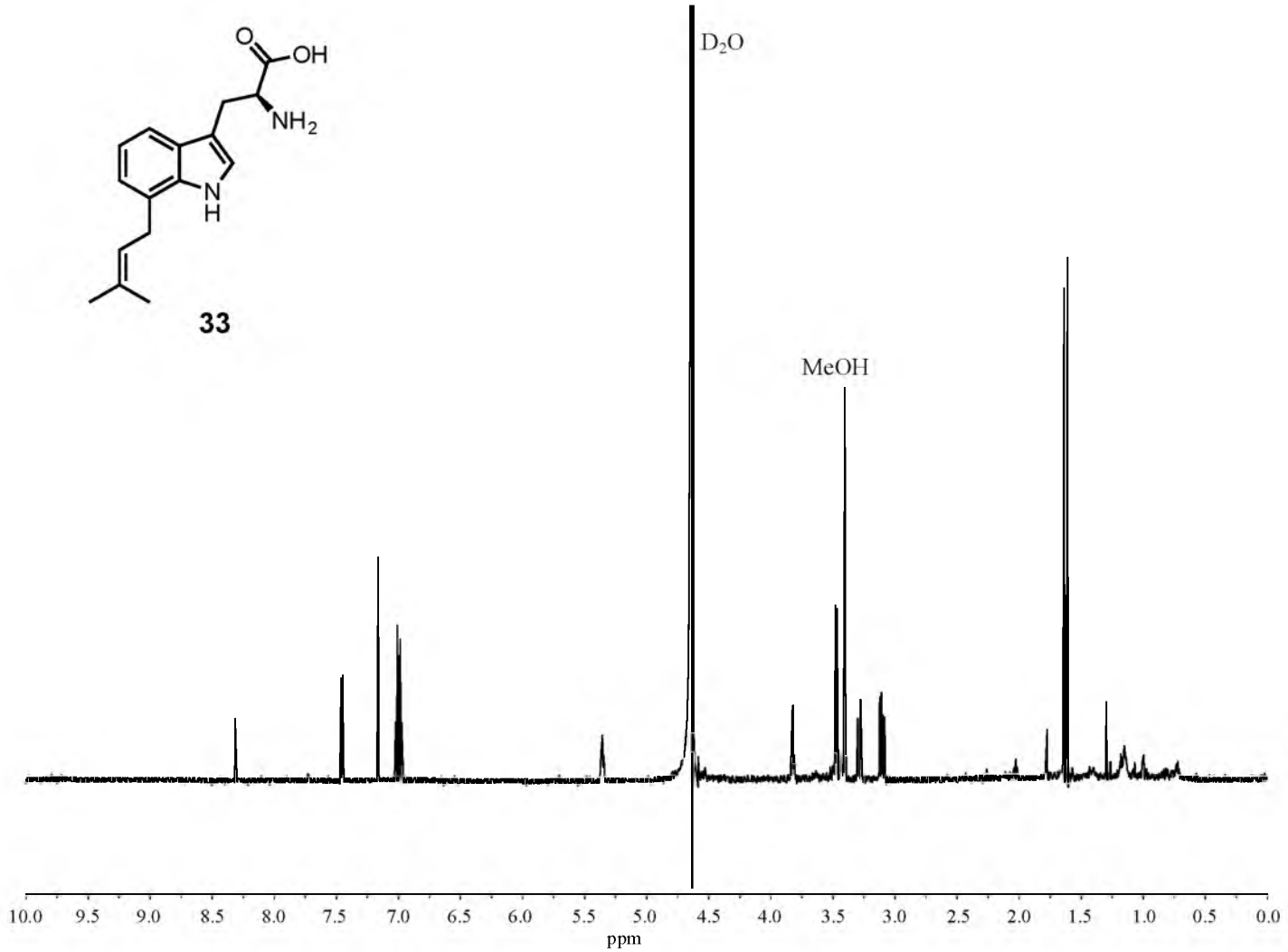
NMR 37. ^{1}H-^{1}H ROESY NMR spectrum of compound 32 in D_2O.



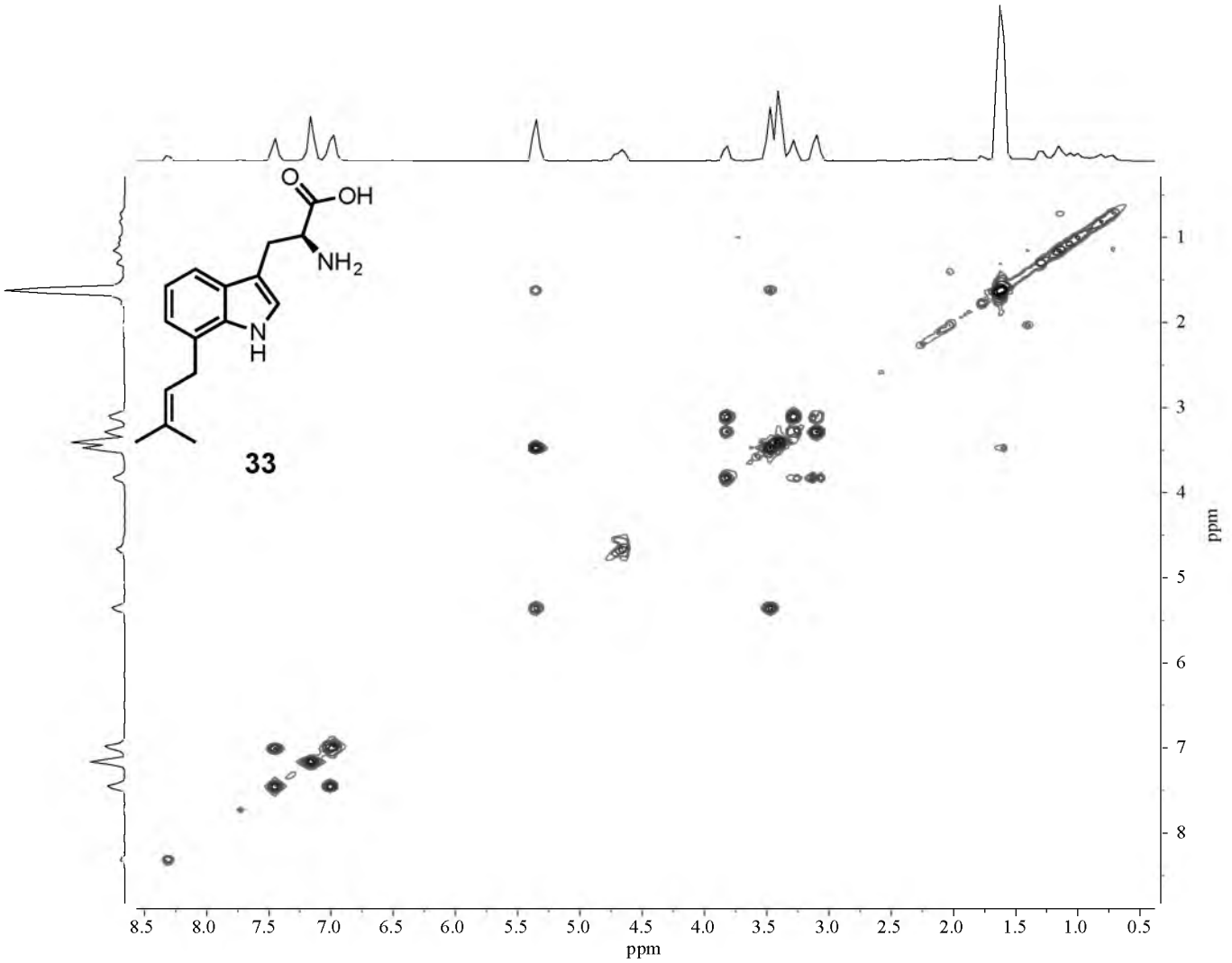
NMR 38. ^{1}H-^{13}C HMQC NMR spectrum of compound 32 in D_2O.



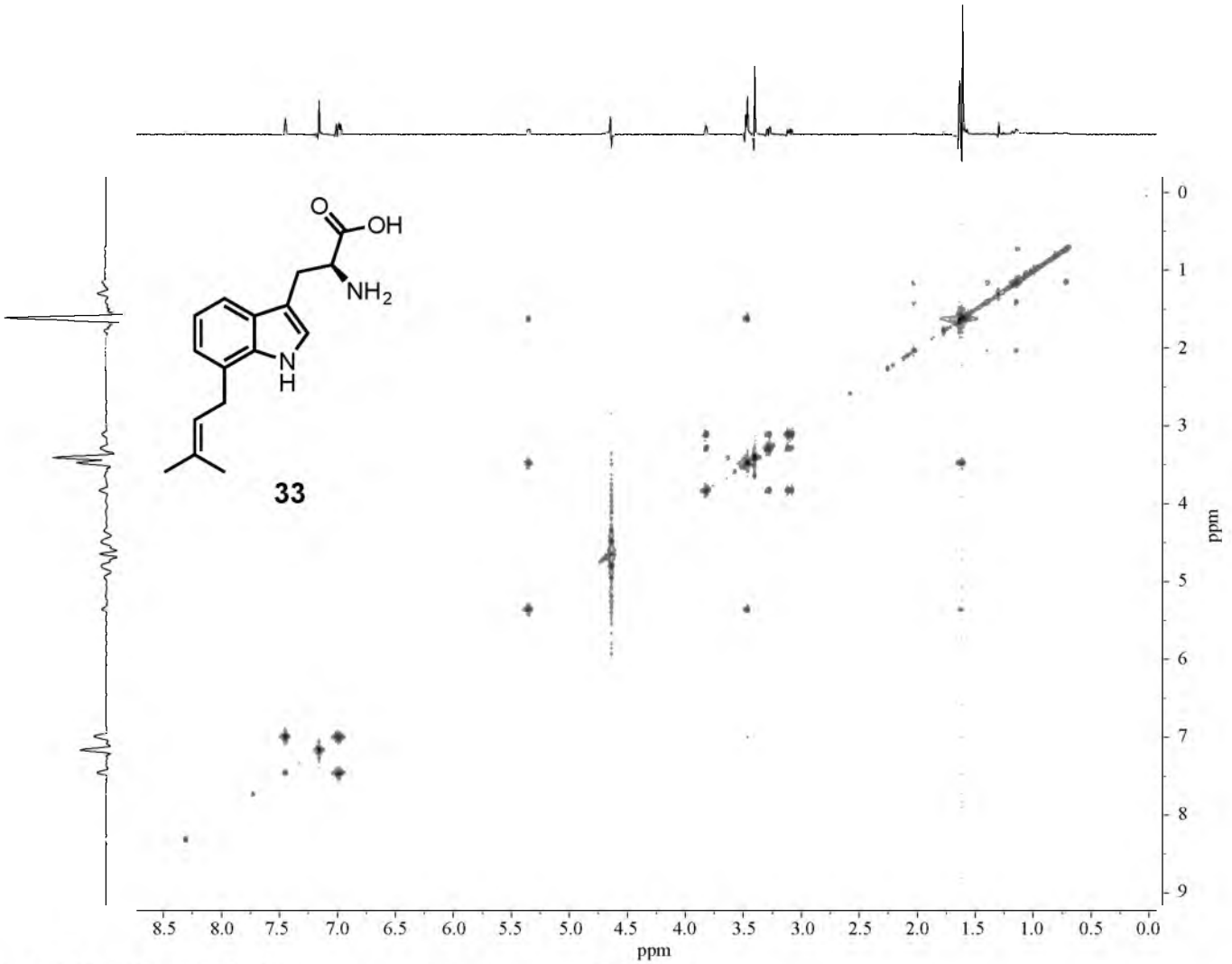
NMR 39. ^{1}H-^{13}C HMBC NMR spectrum of compound 32 in D_2O.



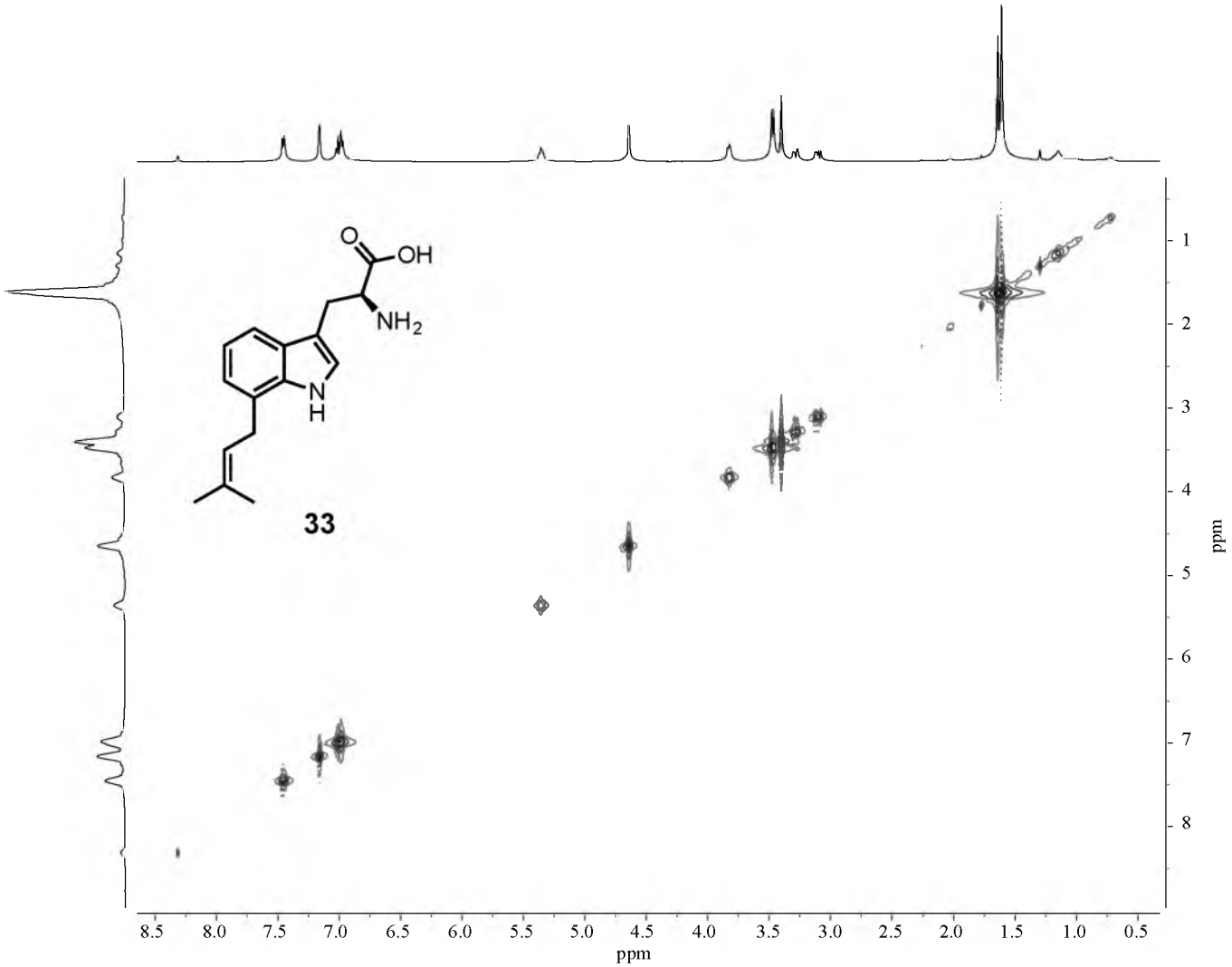
NMR 40. 600 MHz ^{1}H NMR spectrum of compound 33 in D_2O.



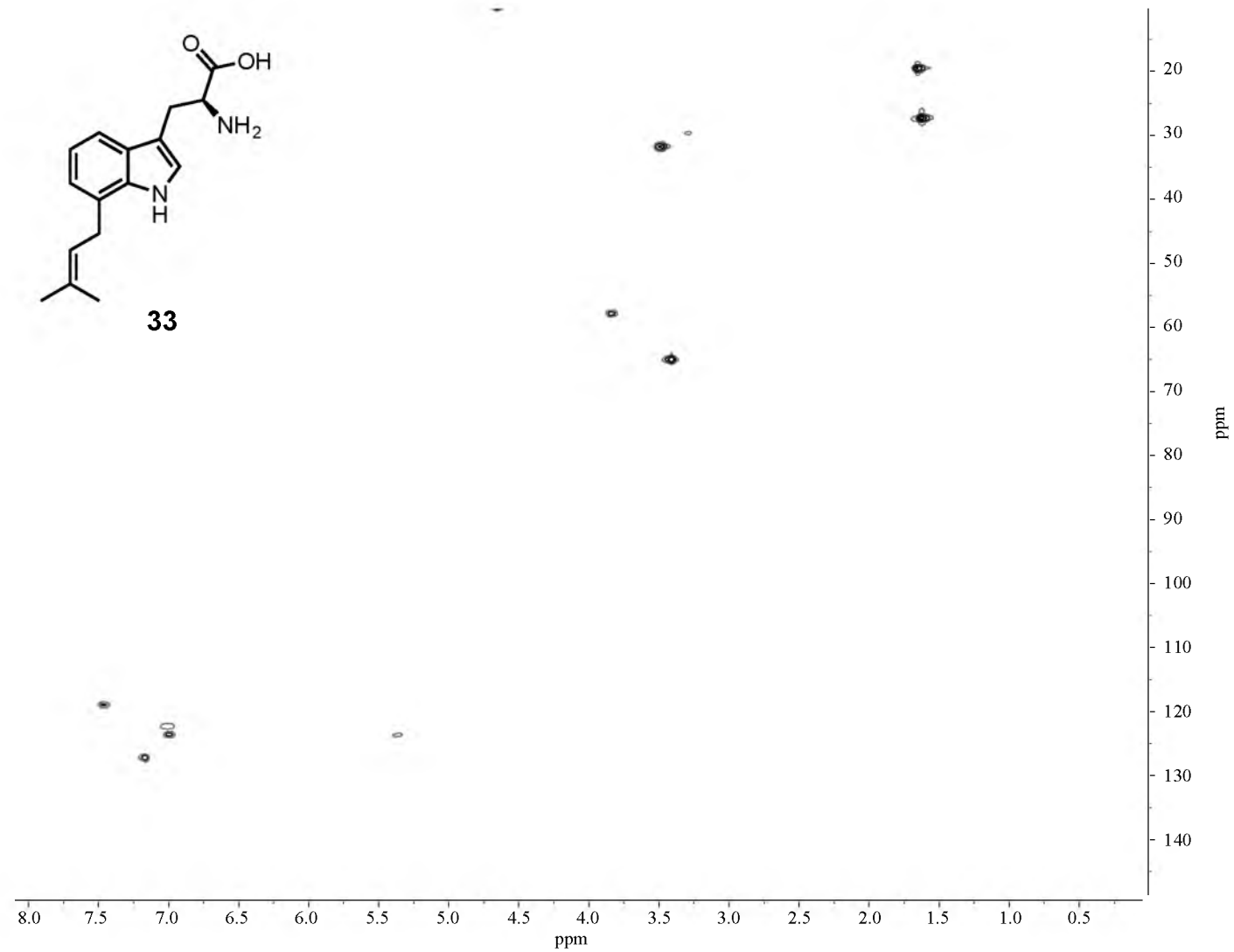
NMR 41. ^{1}H-^{1}H COSY NMR spectrum of compound 33 in D_2O.



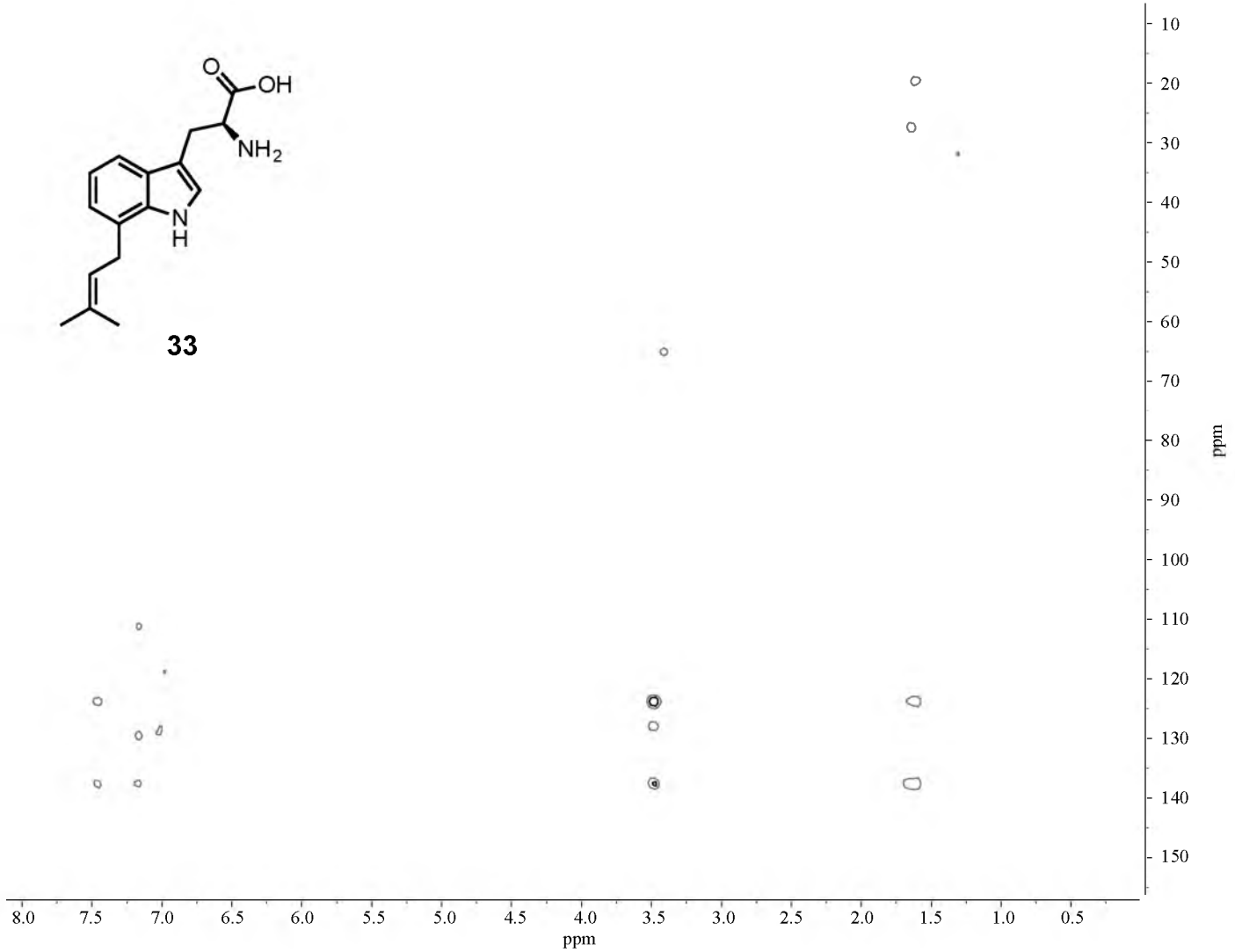
NMR 42. ^{1}H-^{1}H TOCSY NMR spectrum of compound 33 in D_2O.



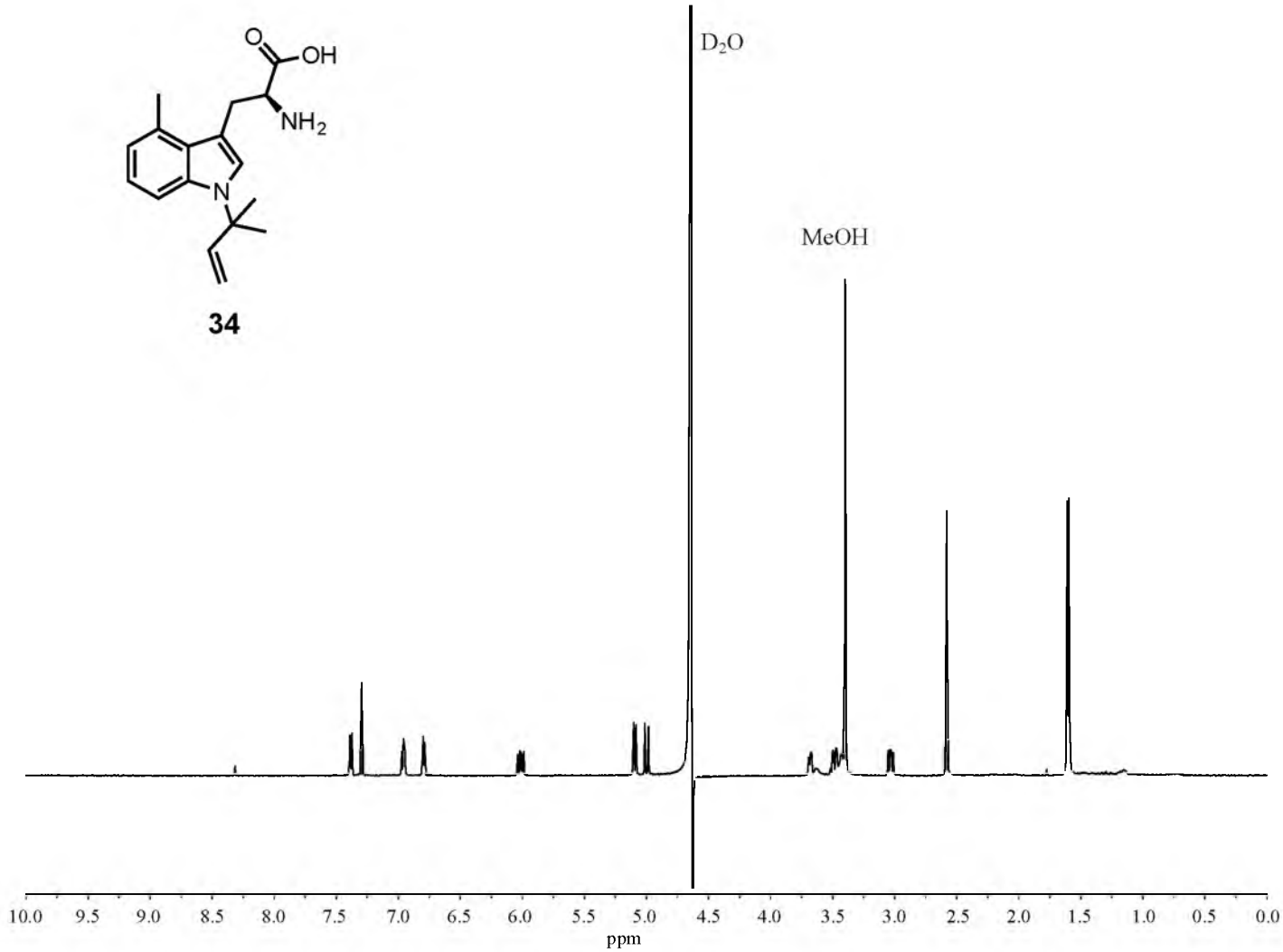
NMR 43. ^{1}H-^{1}H ROESY NMR spectrum of compound 33 in D_2O.



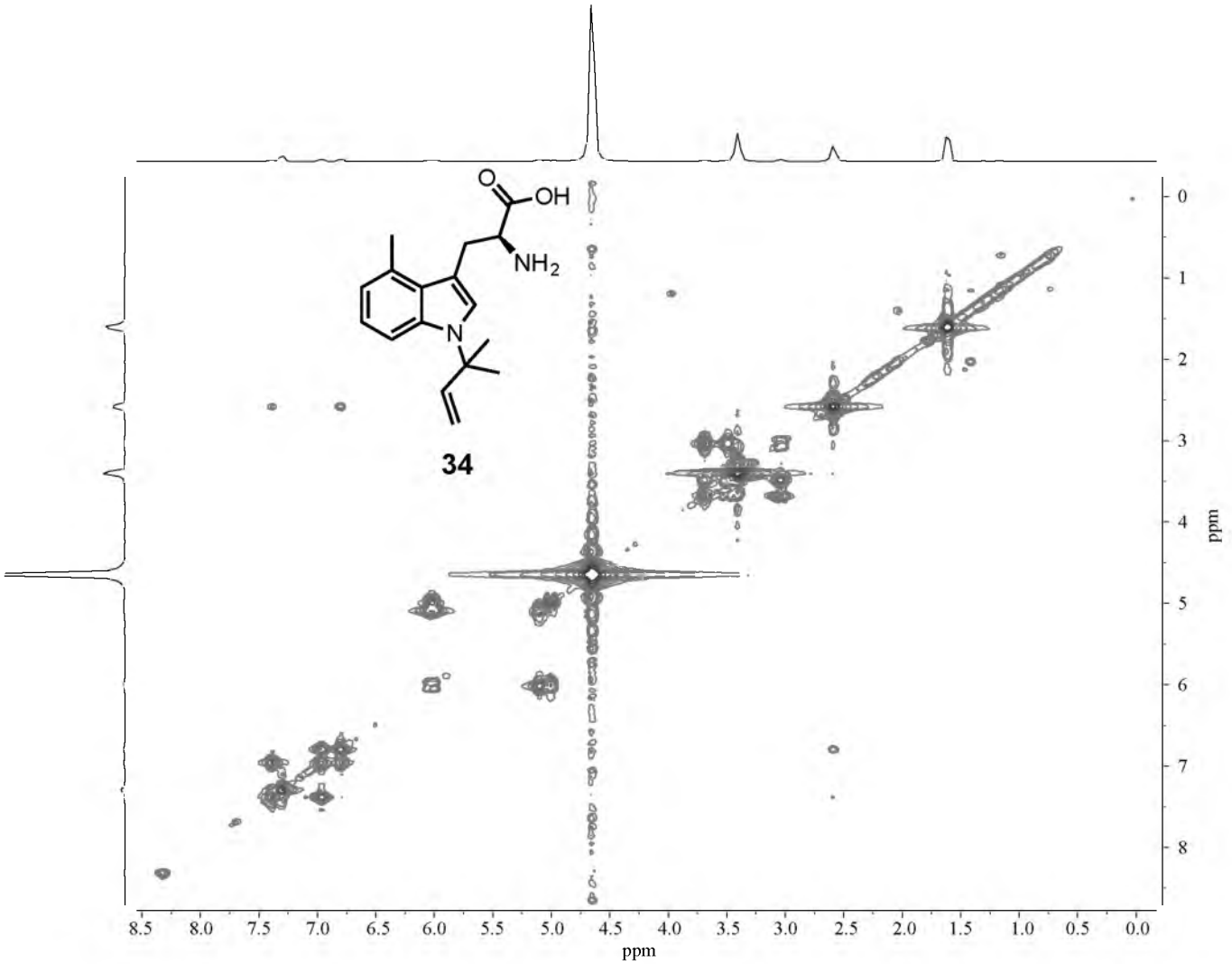
NMR 44. ^{1}H-^{13}C HMQC NMR spectrum of compound 33 in D_2O.



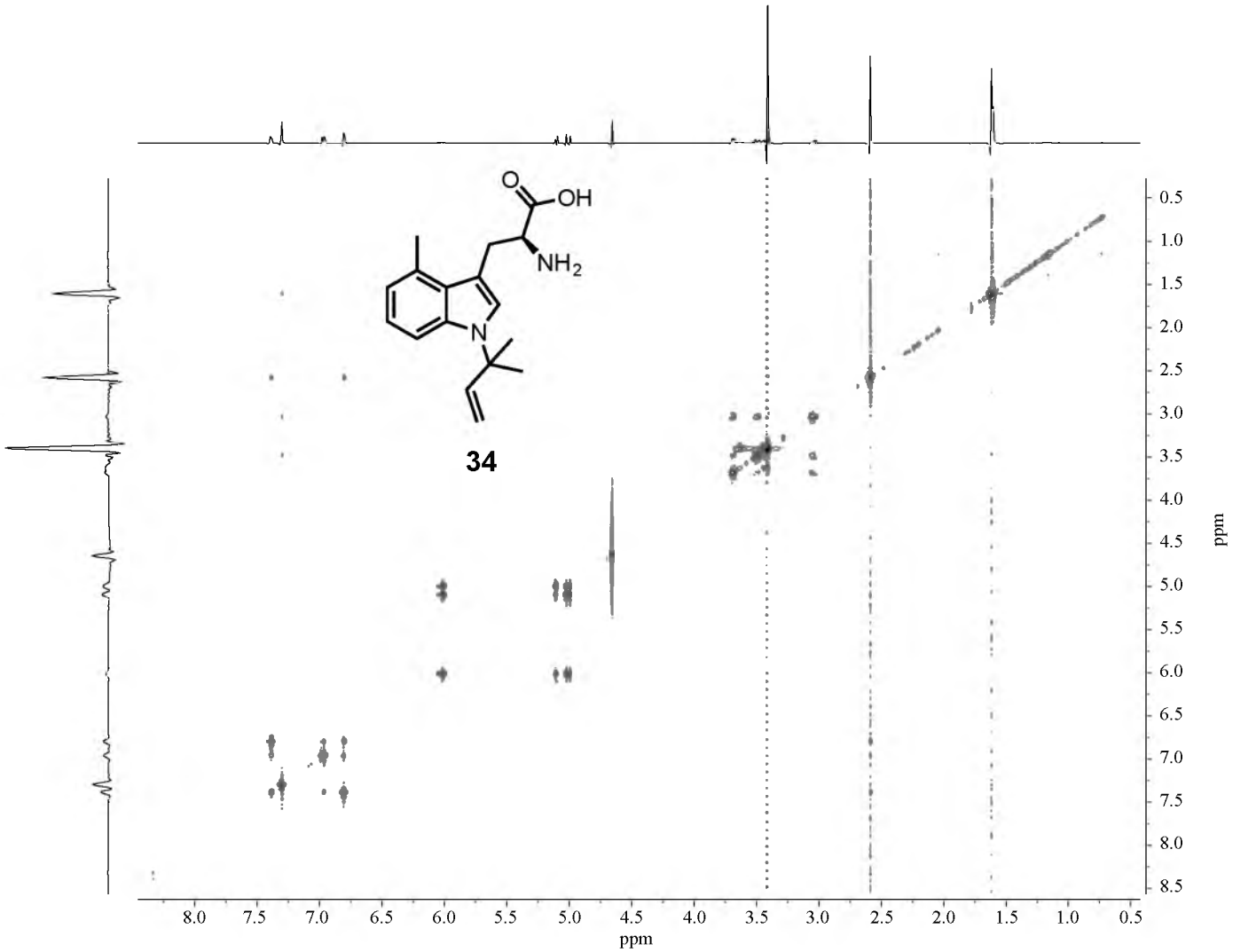
NMR 45. ^{1}H-^{13}C HMBC NMR spectrum of compound 33 in D_2O.



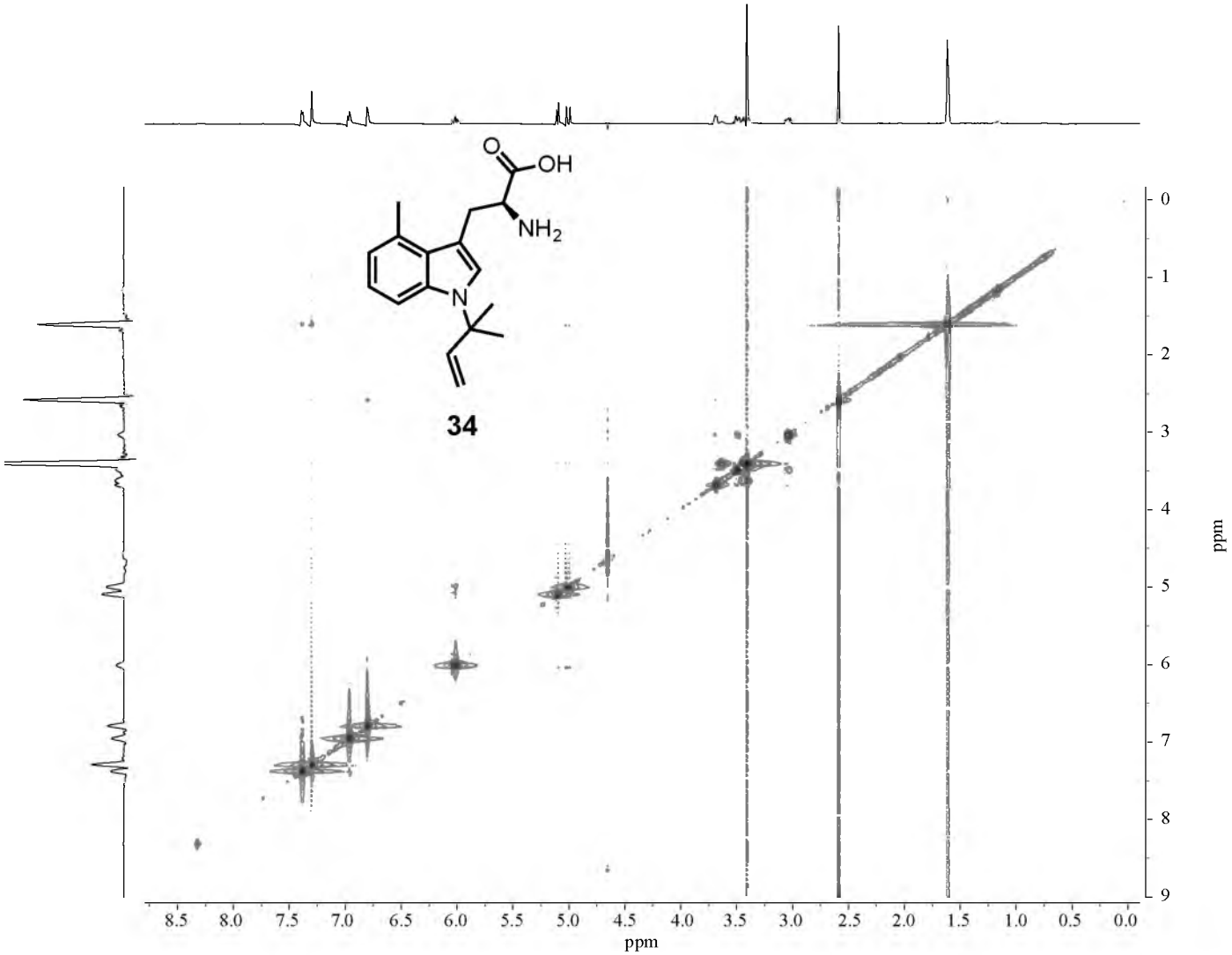
NMR 46. 600 MHz ^{1}H NMR spectrum of compound 34 in D_2O.



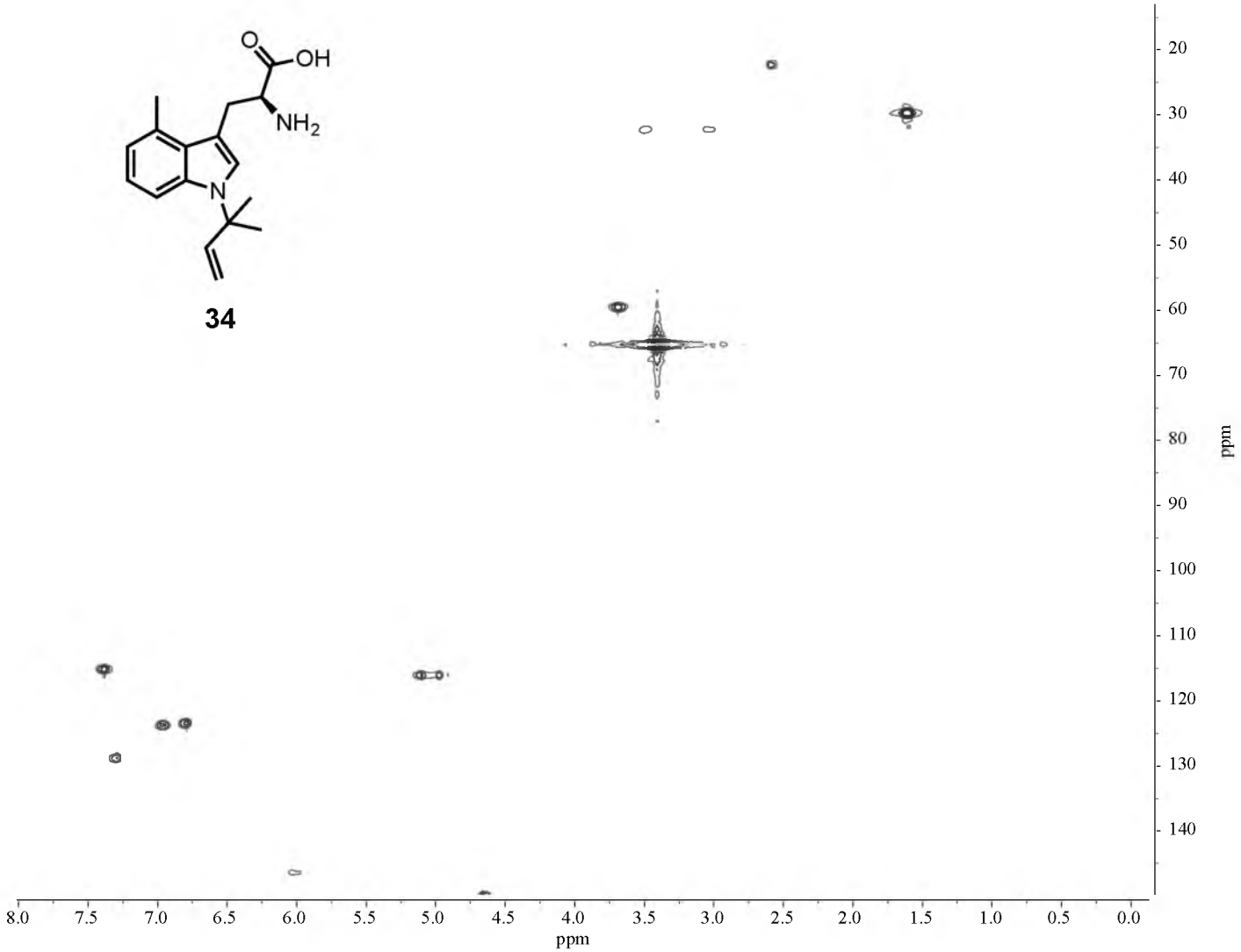
NMR 47. ^{1}H-^{1}H COSY NMR spectrum of compound 34 in D_2O.



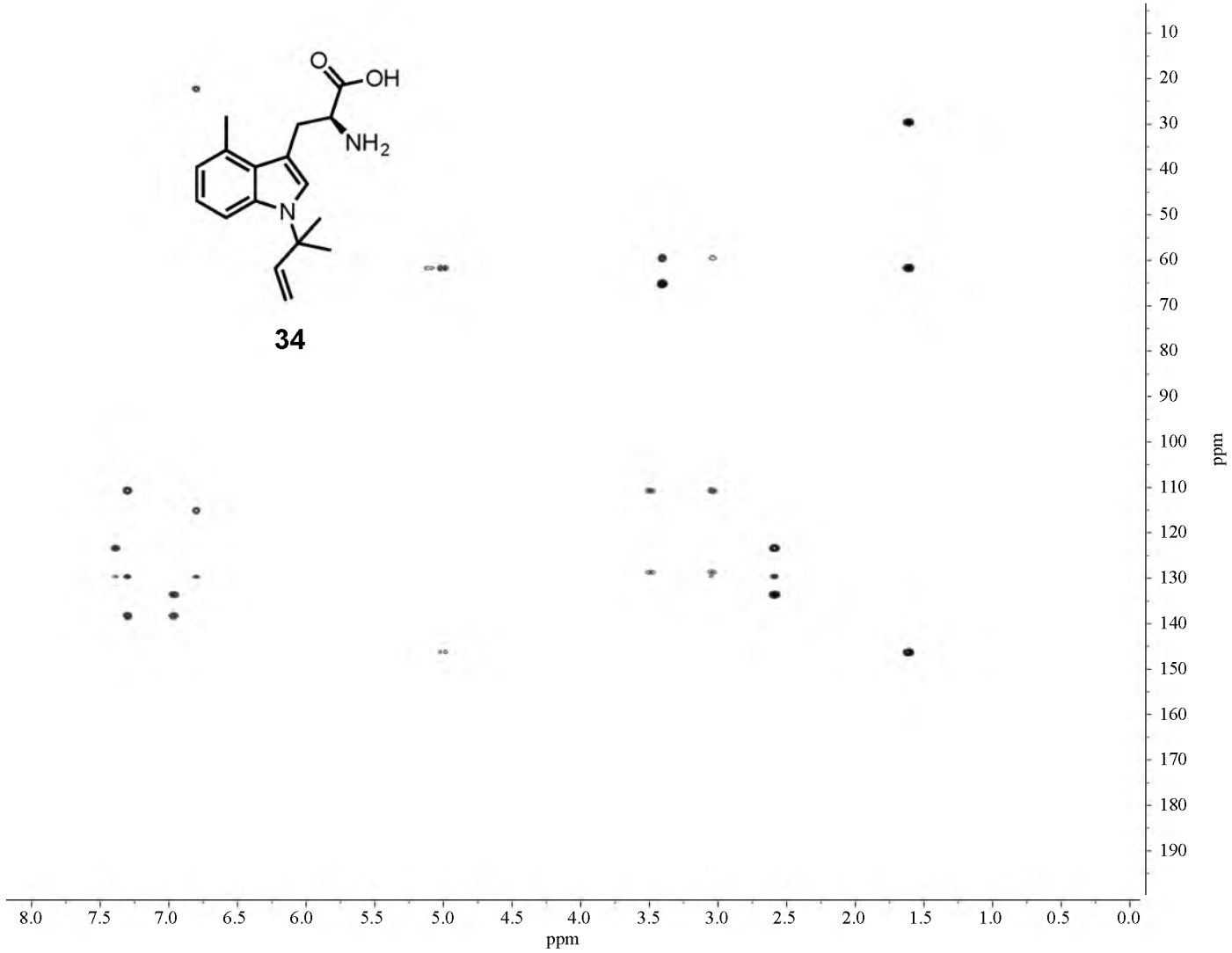
NMR 48. ^{1}H-^{1}H TOCSY NMR spectrum of compound 34 in D_2O.



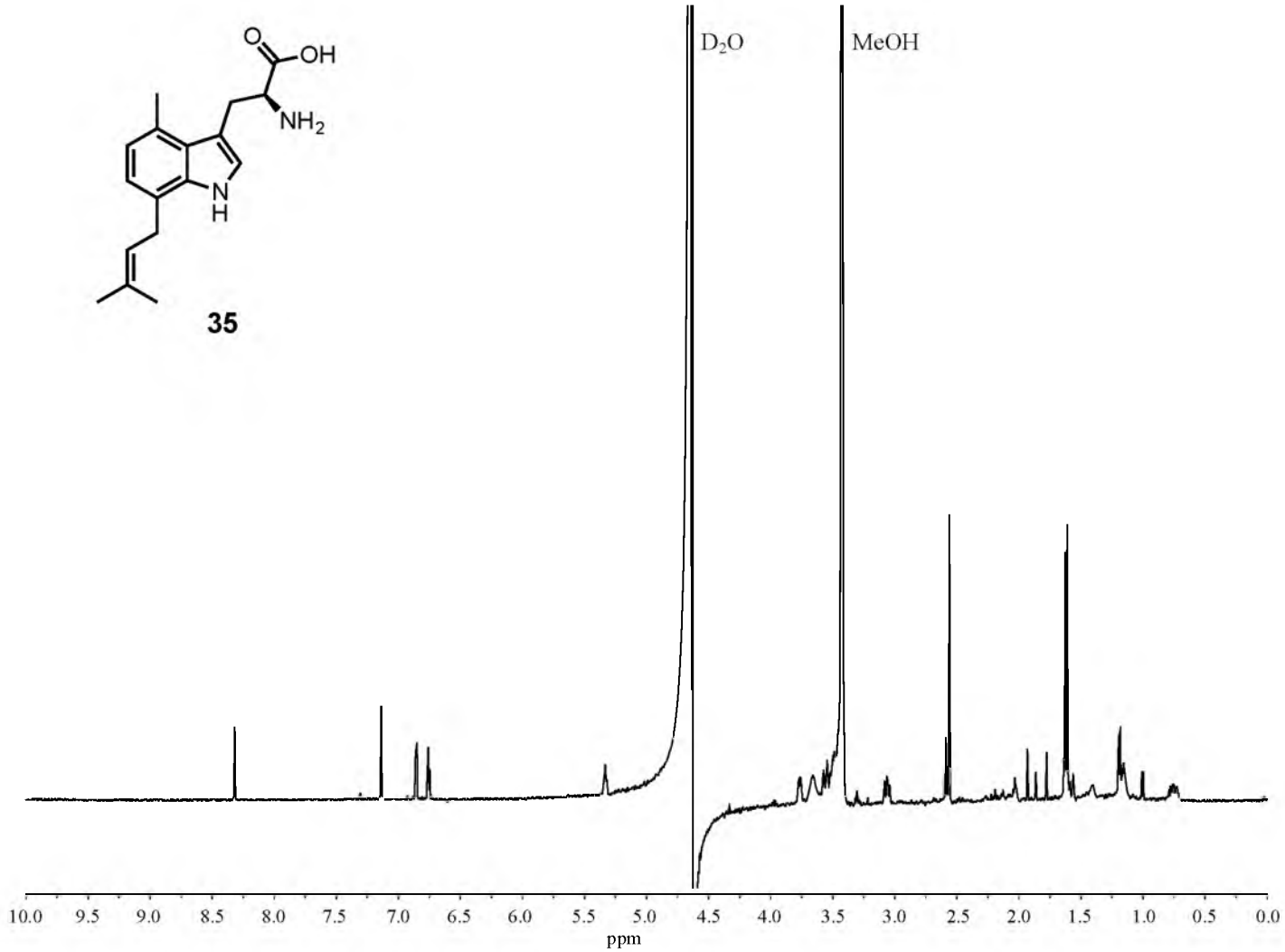
NMR 49. ^{1}H-^{1}H ROESY NMR spectrum of compound 34 in D_2O.



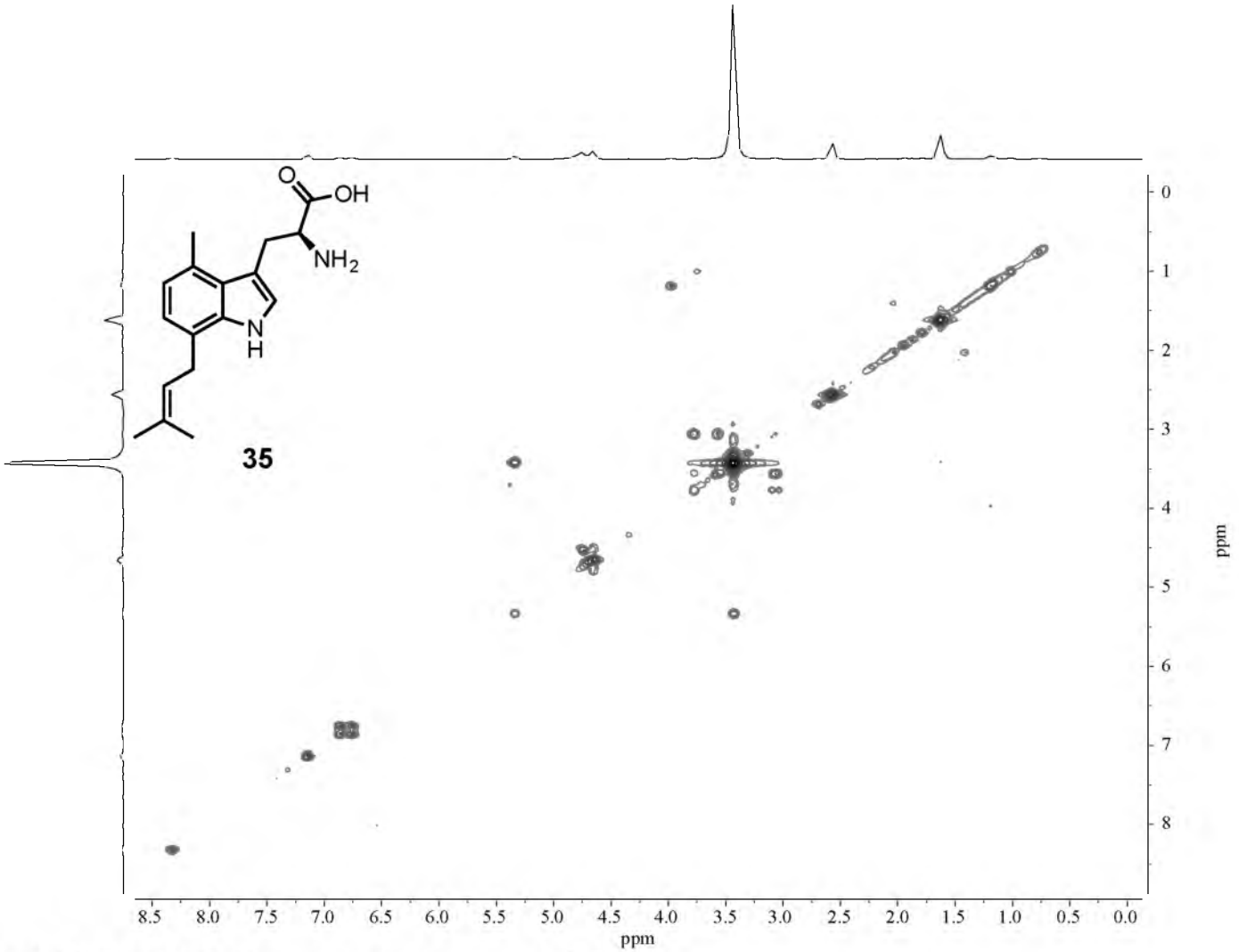
NMR 50. ^{1}H-^{13}C HSQC NMR spectrum of compound 34 in D_2O.



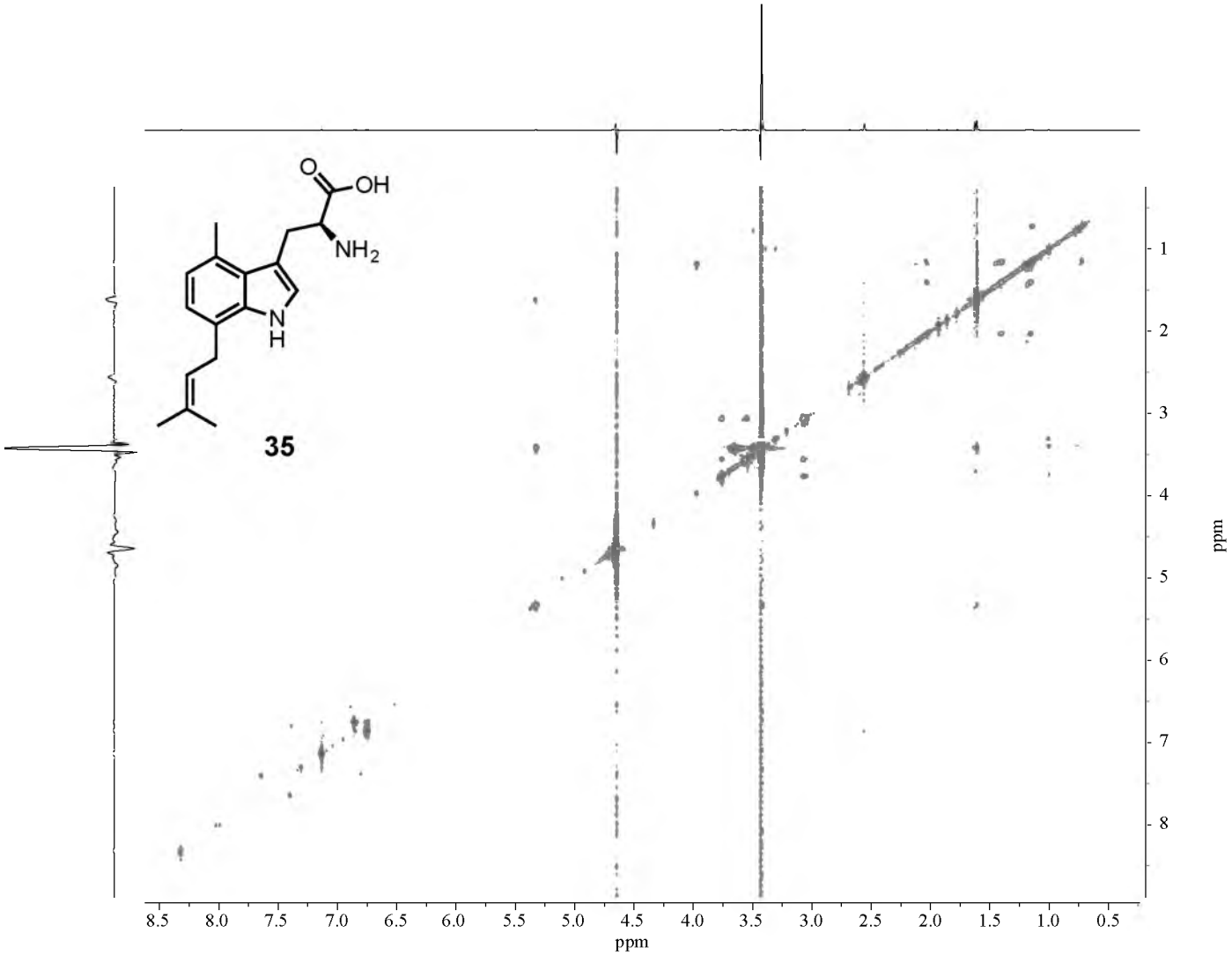
NMR 51. ^{1}H-^{13}C HMBC NMR spectrum of compound 34 in D_2O.



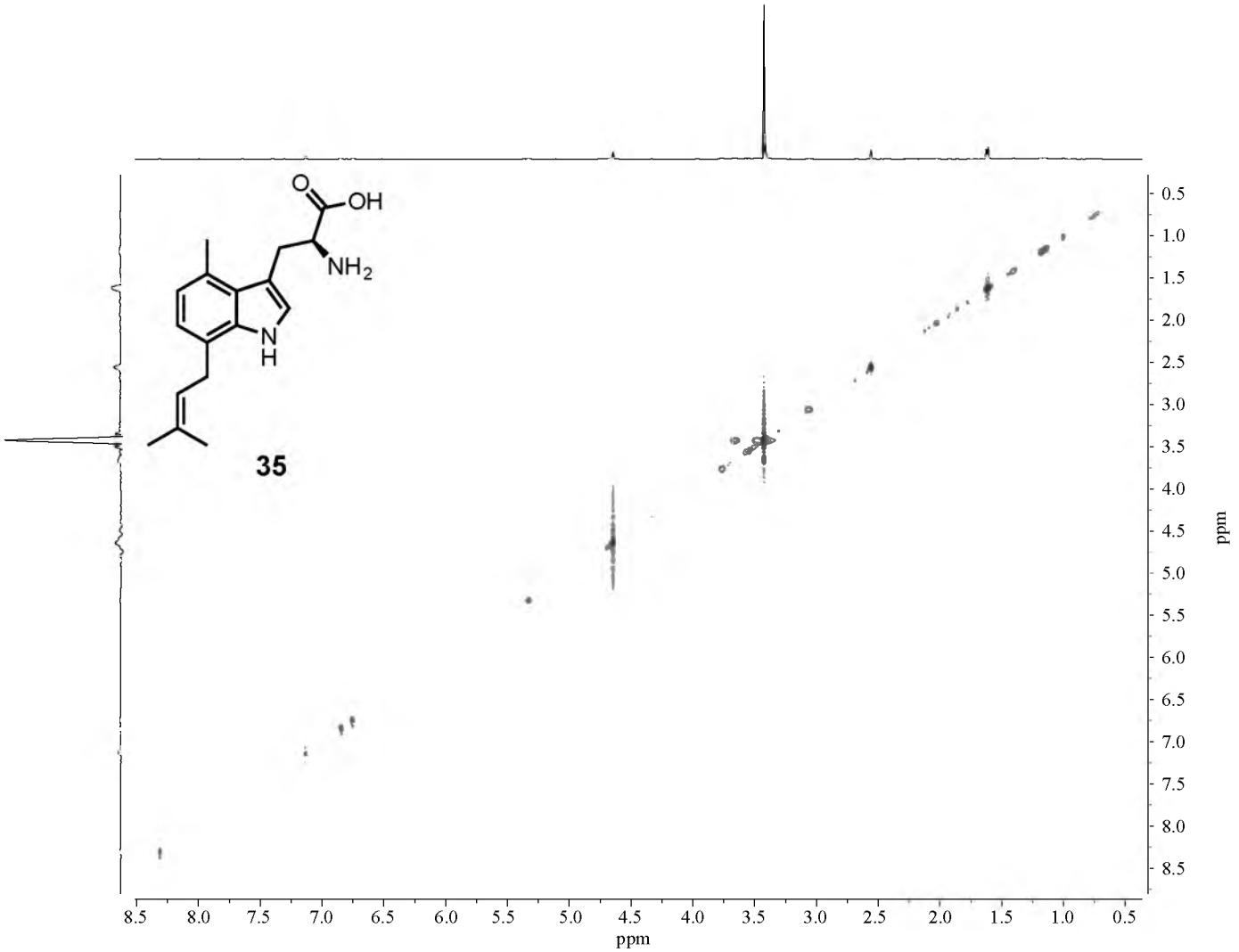
NMR 52. 600 MHz ^{1}H NMR spectrum of compound 35 in D_2O.



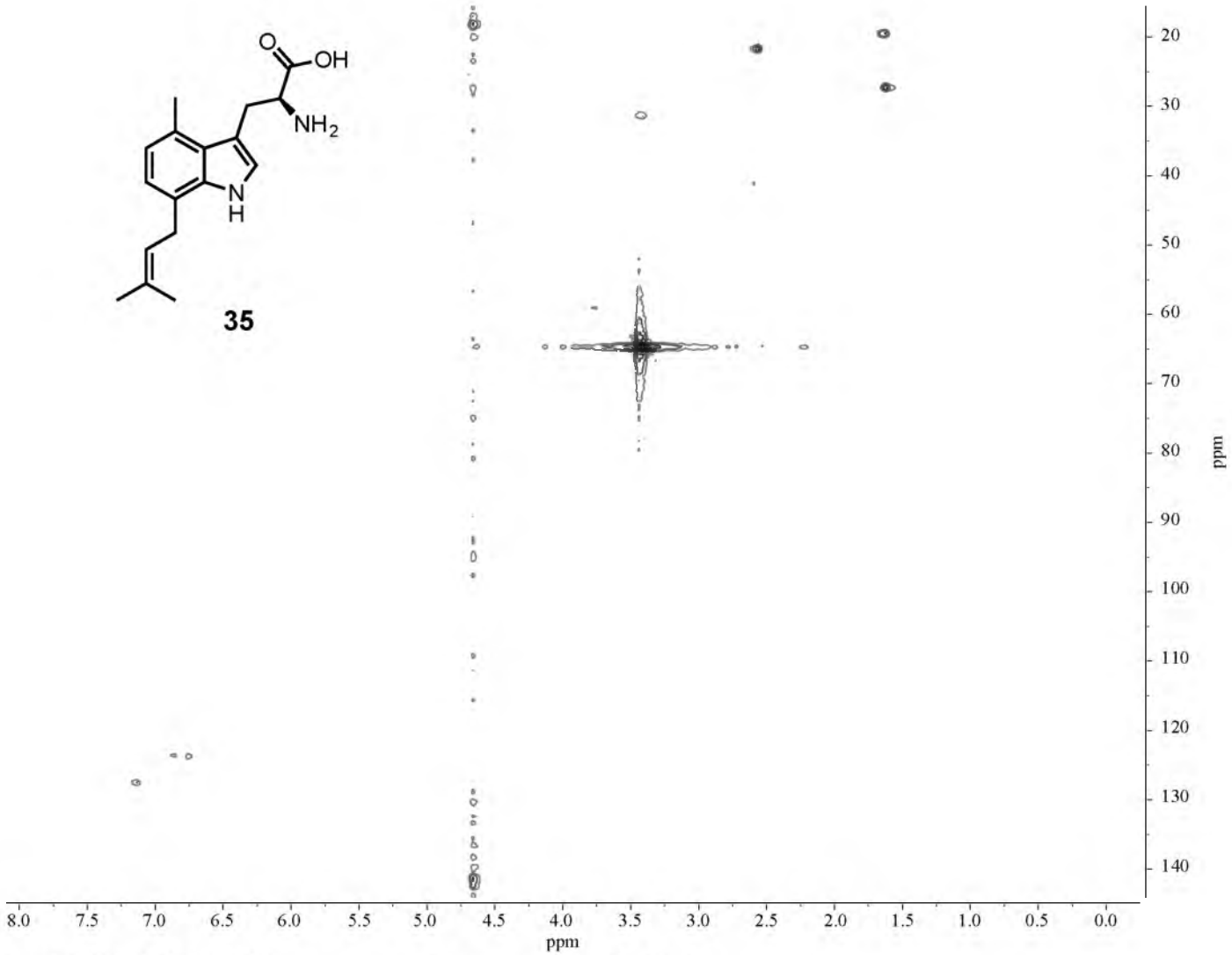
NMR 53. ^{1}H-^{1}H COSY NMR spectrum of compound 35 in D_2O.



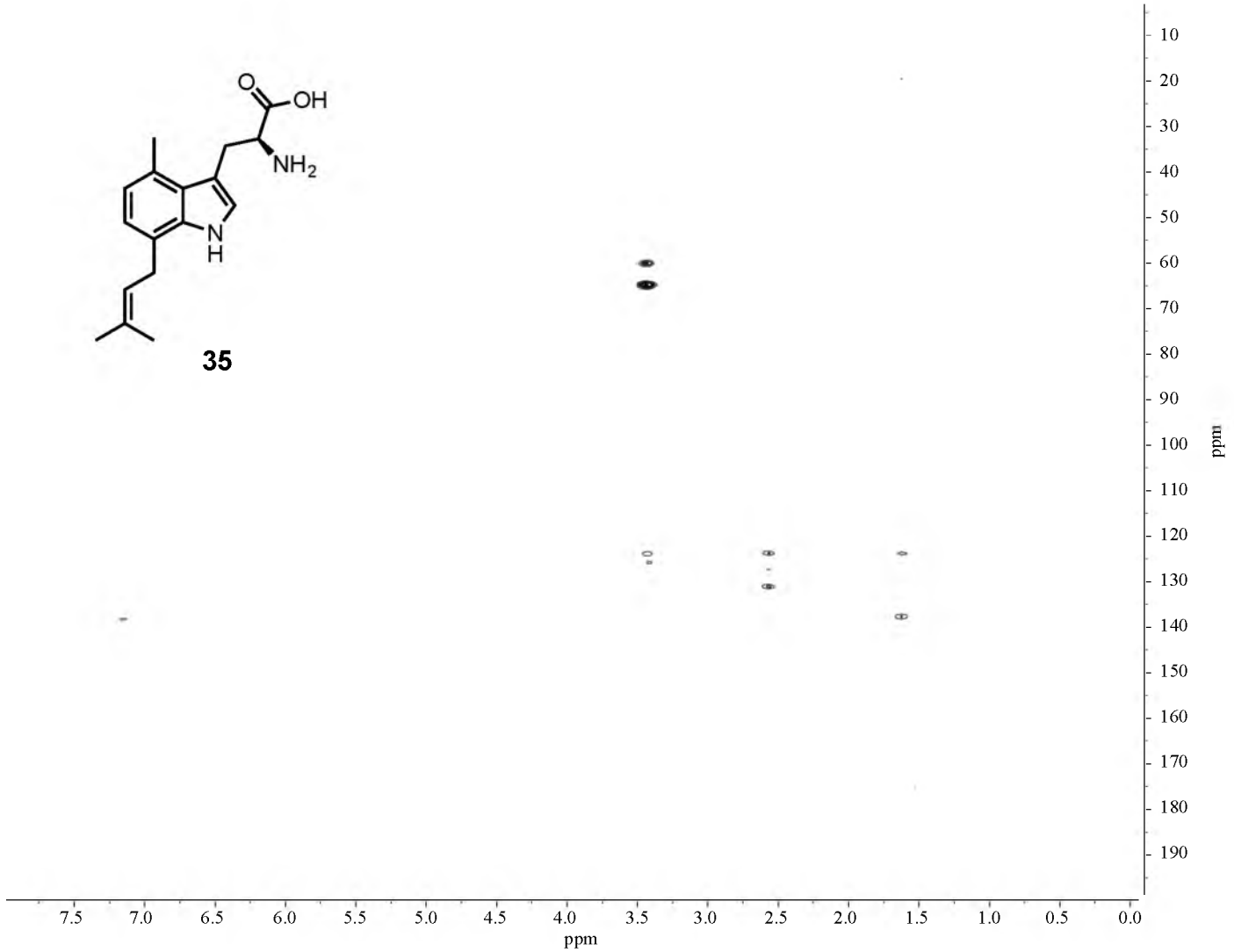
NMR 54. ^{1}H-^{1}H TOCSY NMR spectrum of compound 35 in D_2O.



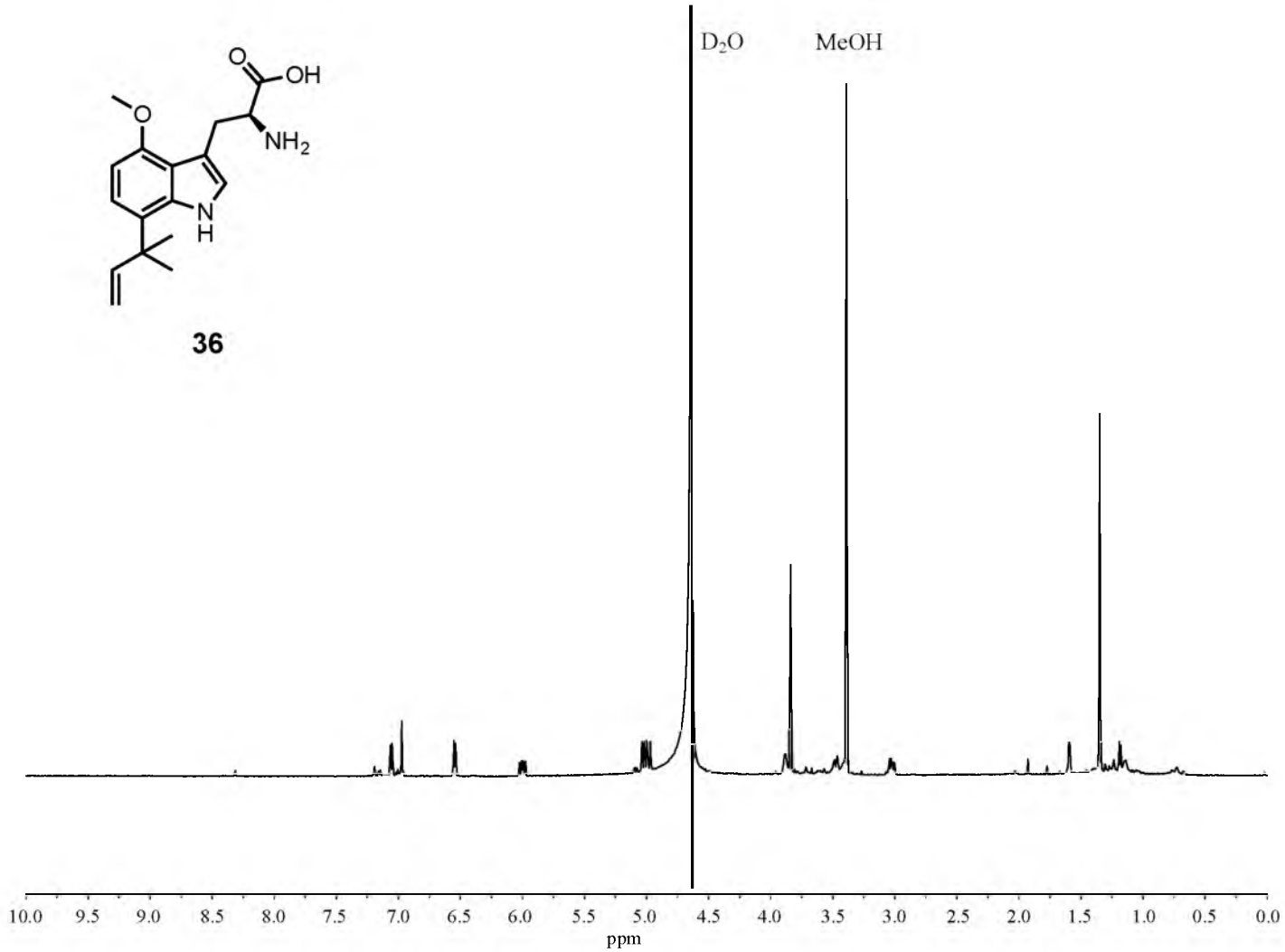
NMR 55. ^{1}H-^{1}H ROESY NMR spectrum of compound 35 in D_2O.



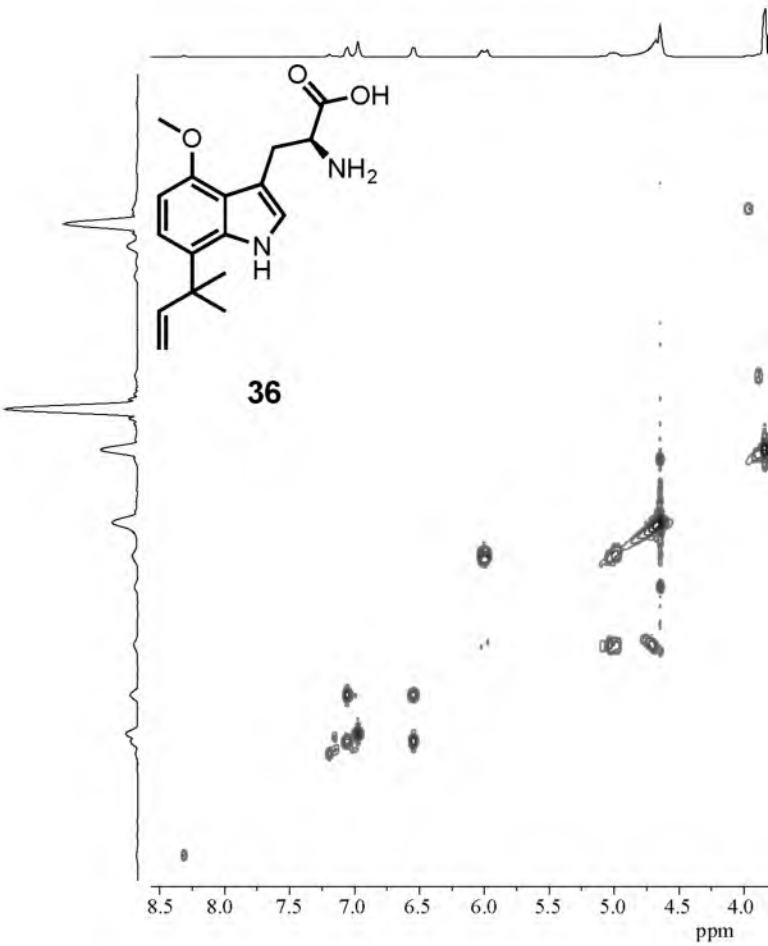
NMR 56. ^{1}H-^{13}C HMQC NMR spectrum of compound 35 in D_2O.



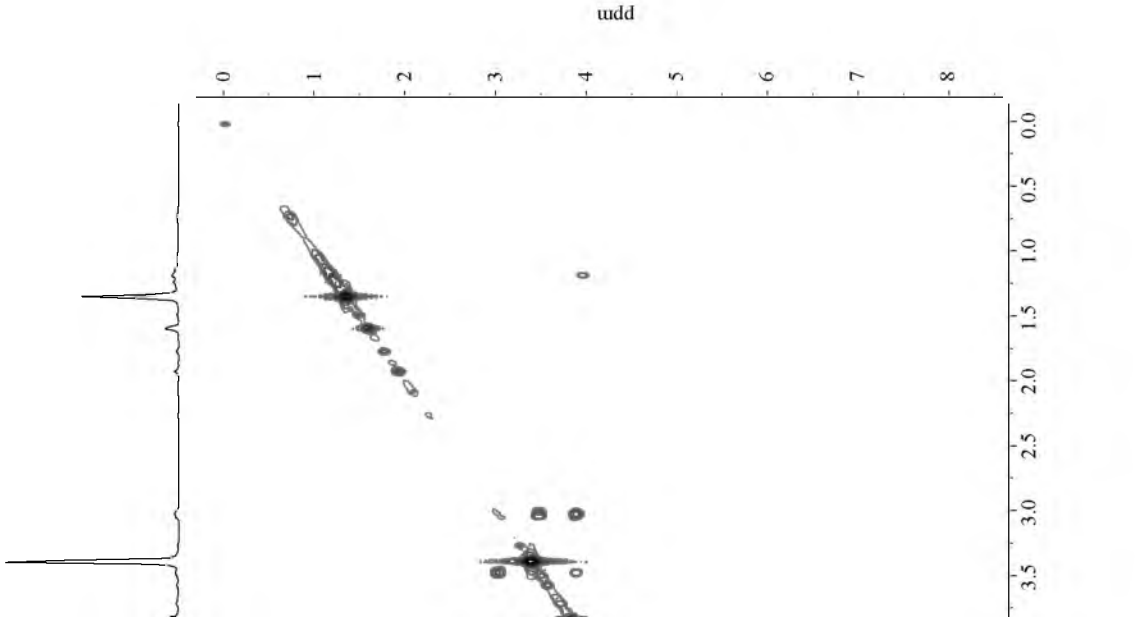
NMR 57. ^{1}H-^{13}C HMBC NMR spectrum of compound 35 in D_2O.

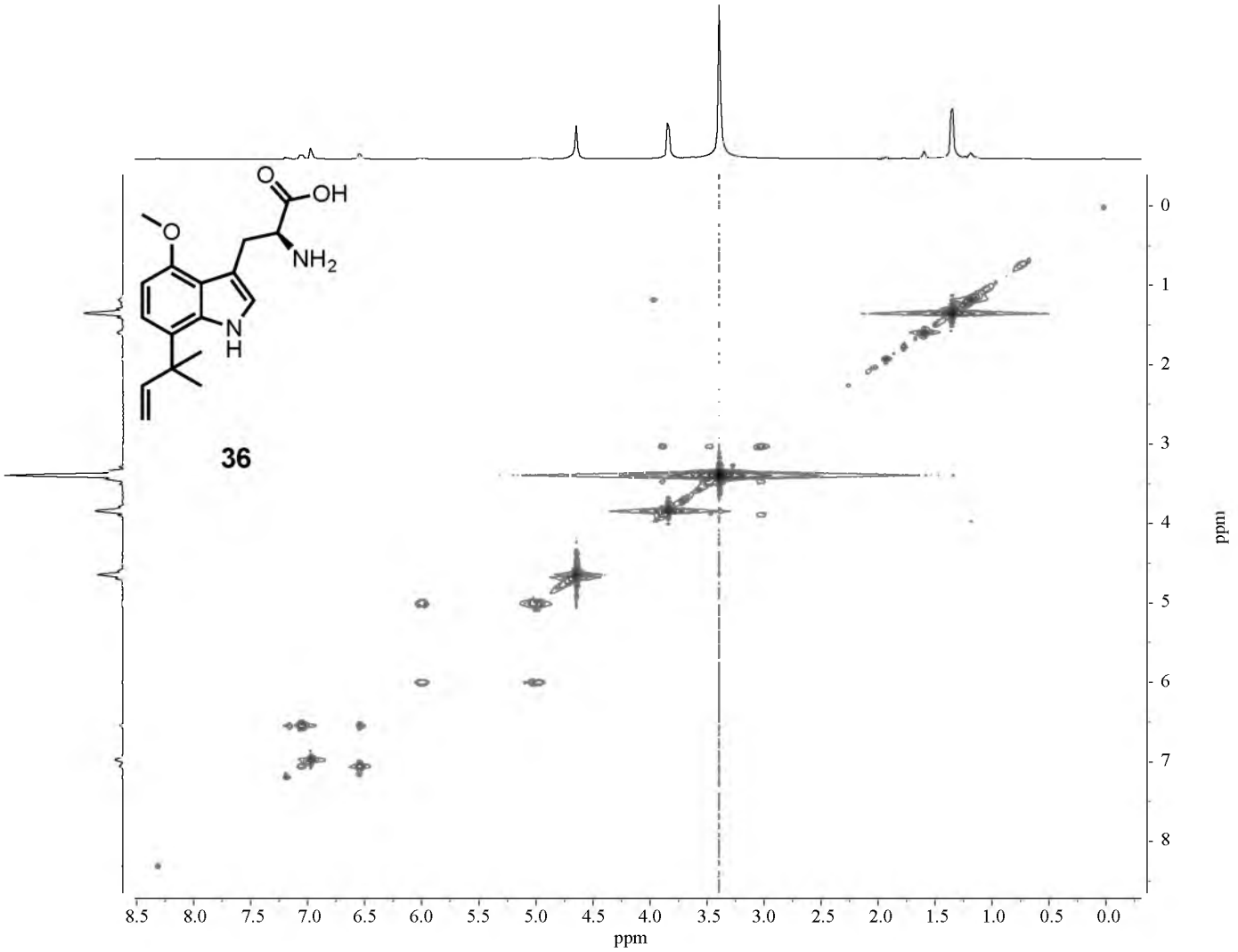


NMR 58. 600 MHz ^{1}H NMR spectrum of compound 36 in D_2O.

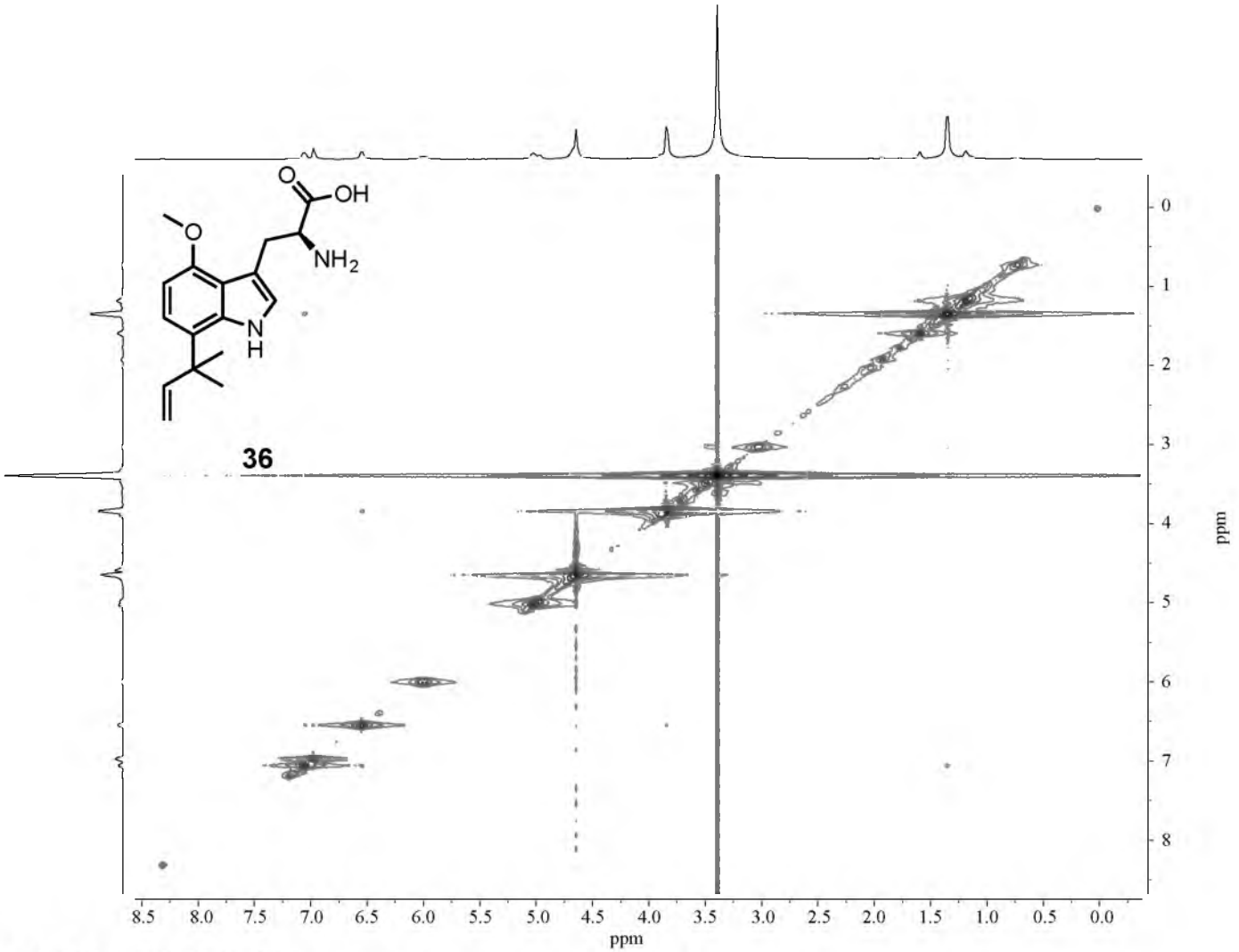


NMR 59. ^{1}H-^{1}H COSY NMR spectrum of compound 36 in D_2O.

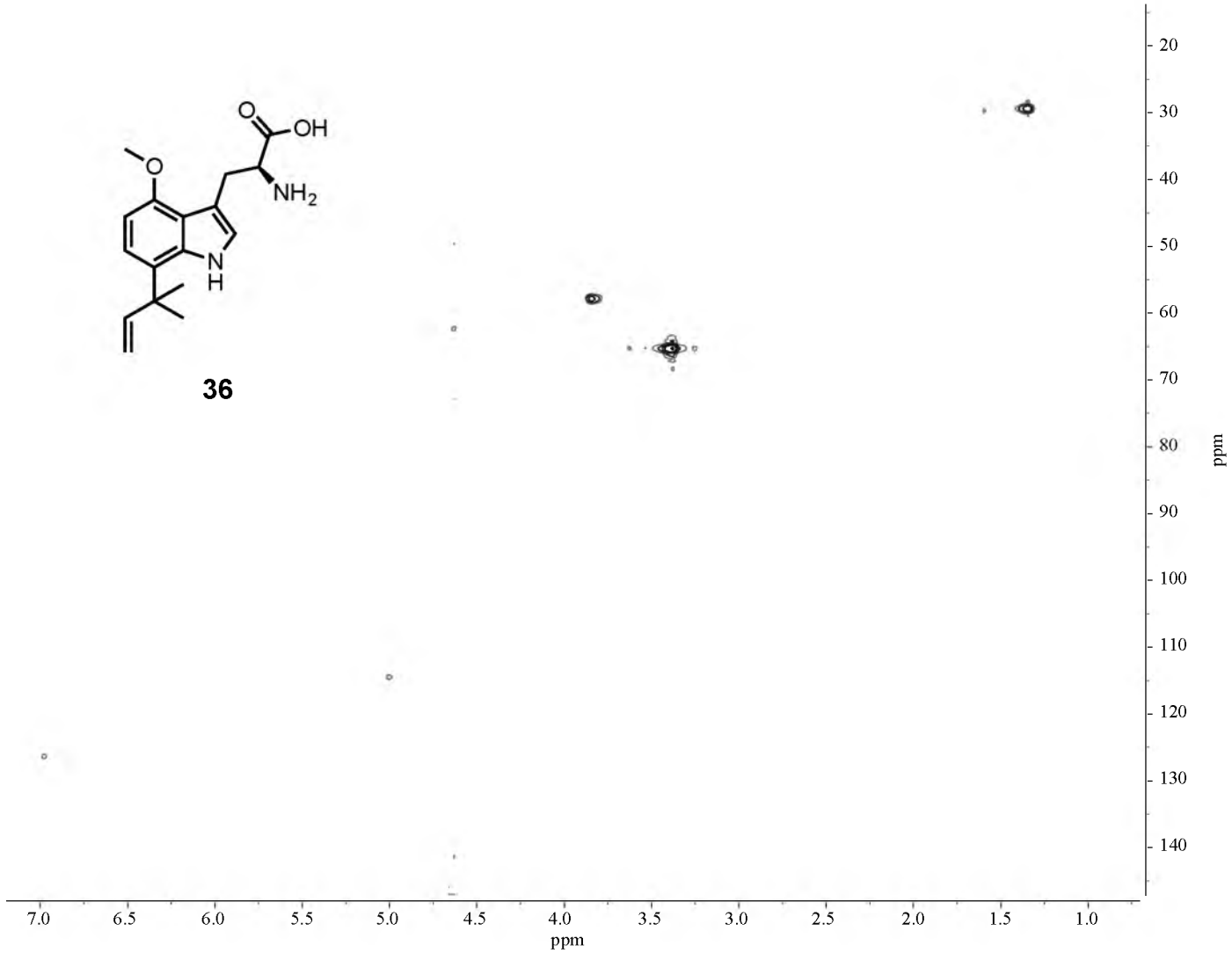

ppm
0
1
2
3
4
5
6
7
8
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5



NMR 60. ^{1}H-^{1}H TOCSY NMR spectrum of compound 36 in D_2O.

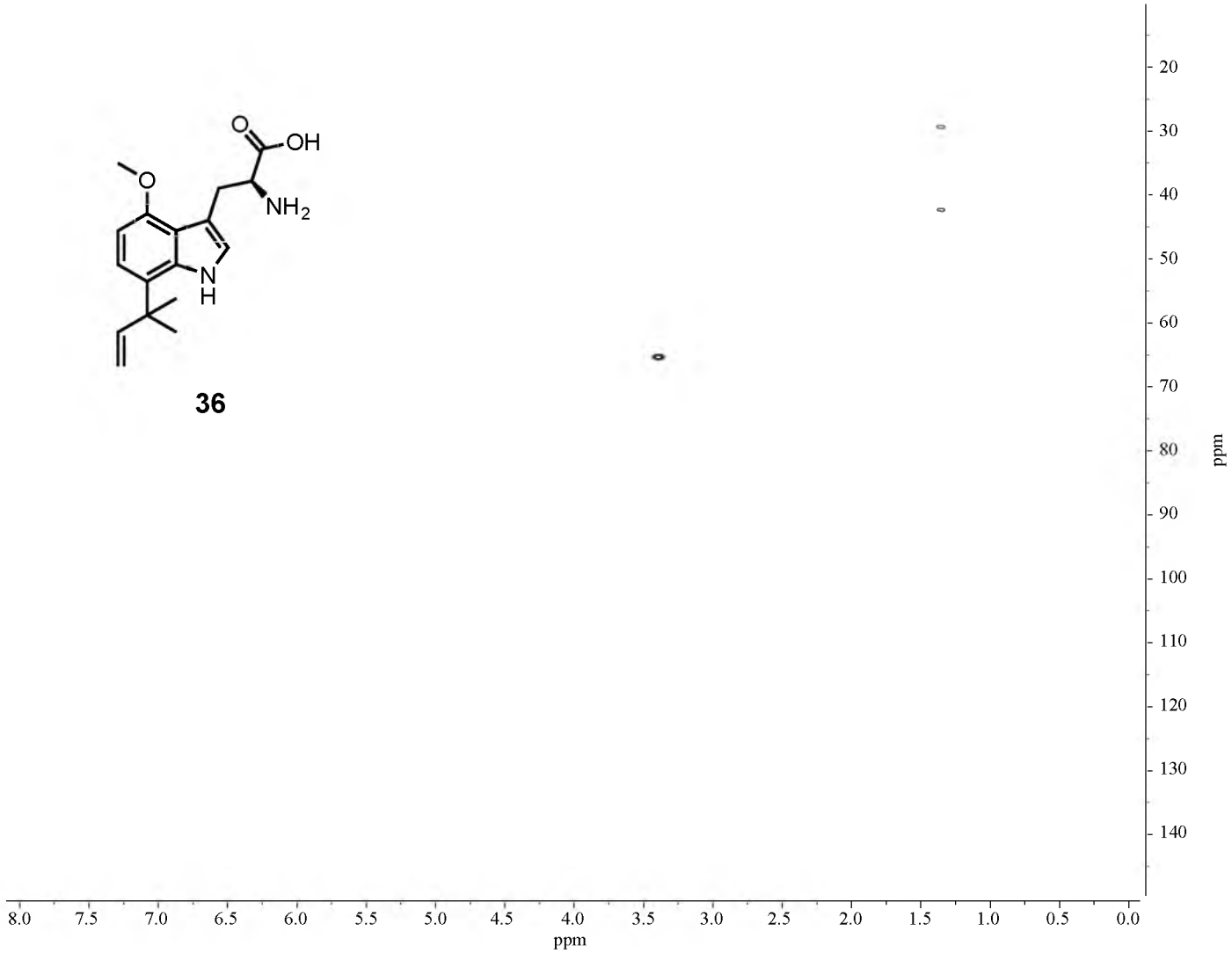


NMR 61. ^{1}H-^{1}H ROESY NMR spectrum of compound 36 in D_2O.

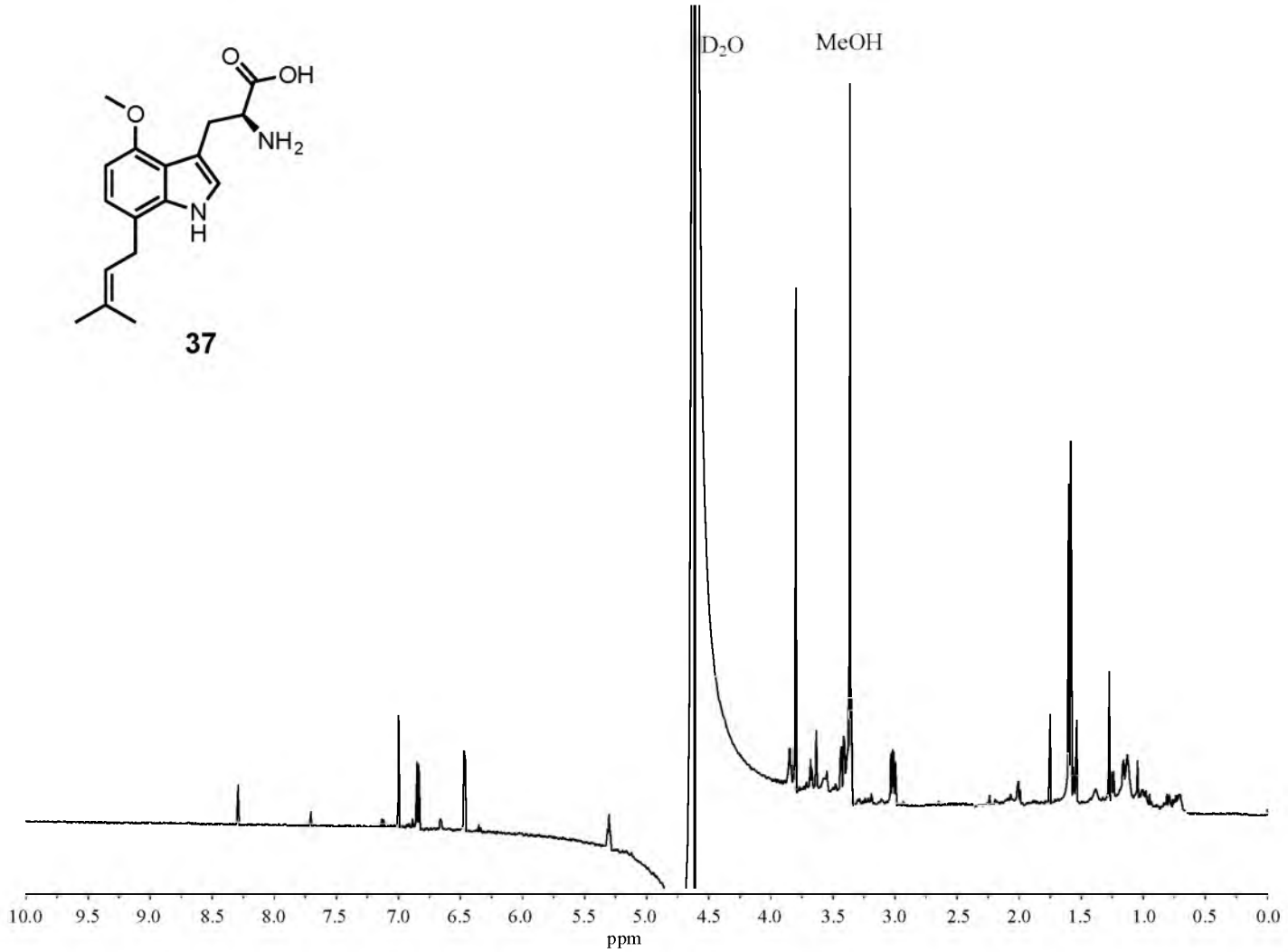


36

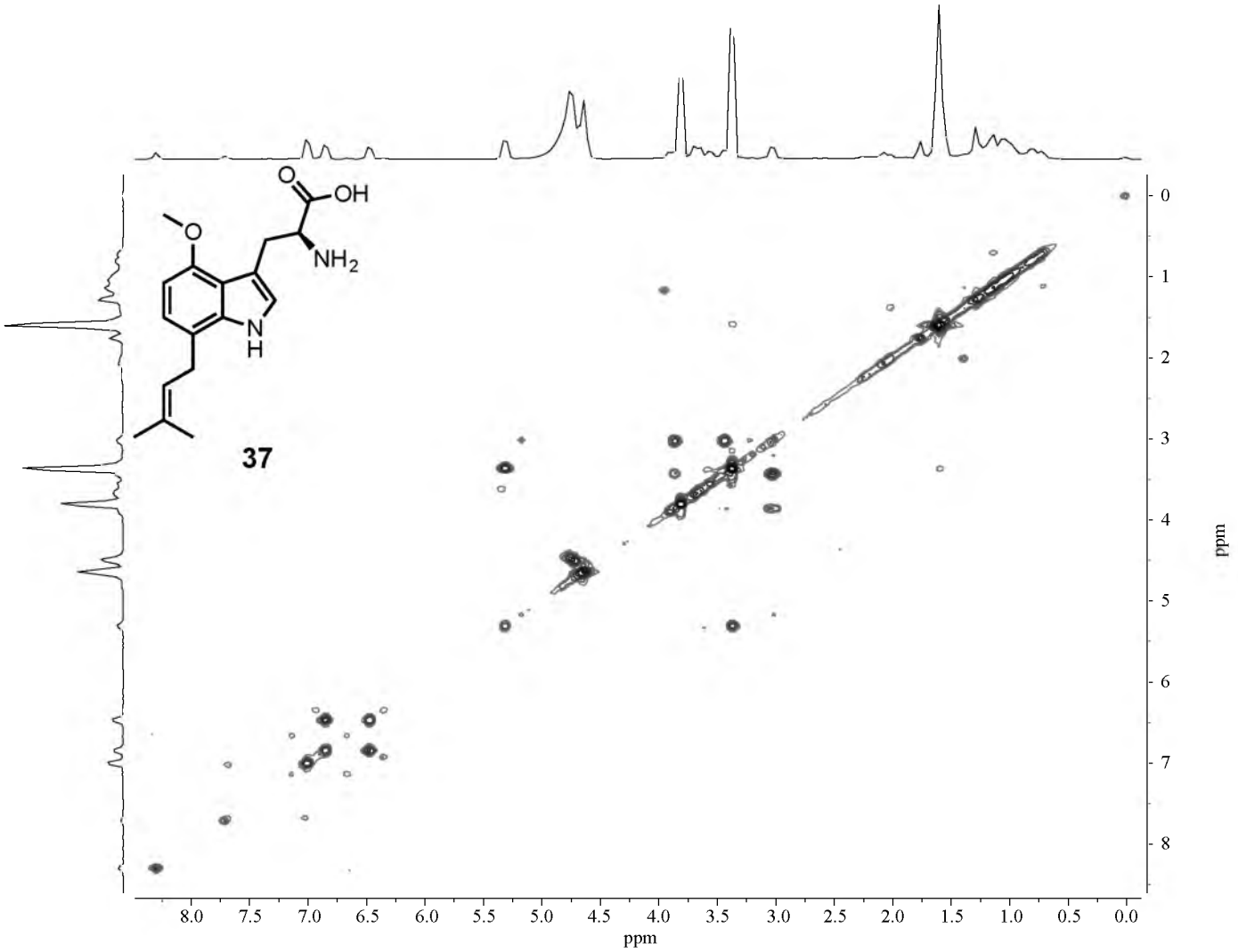
NMR 62. ^{1}H-^{13}C HMQC NMR spectrum of compound 36 in D_2O.



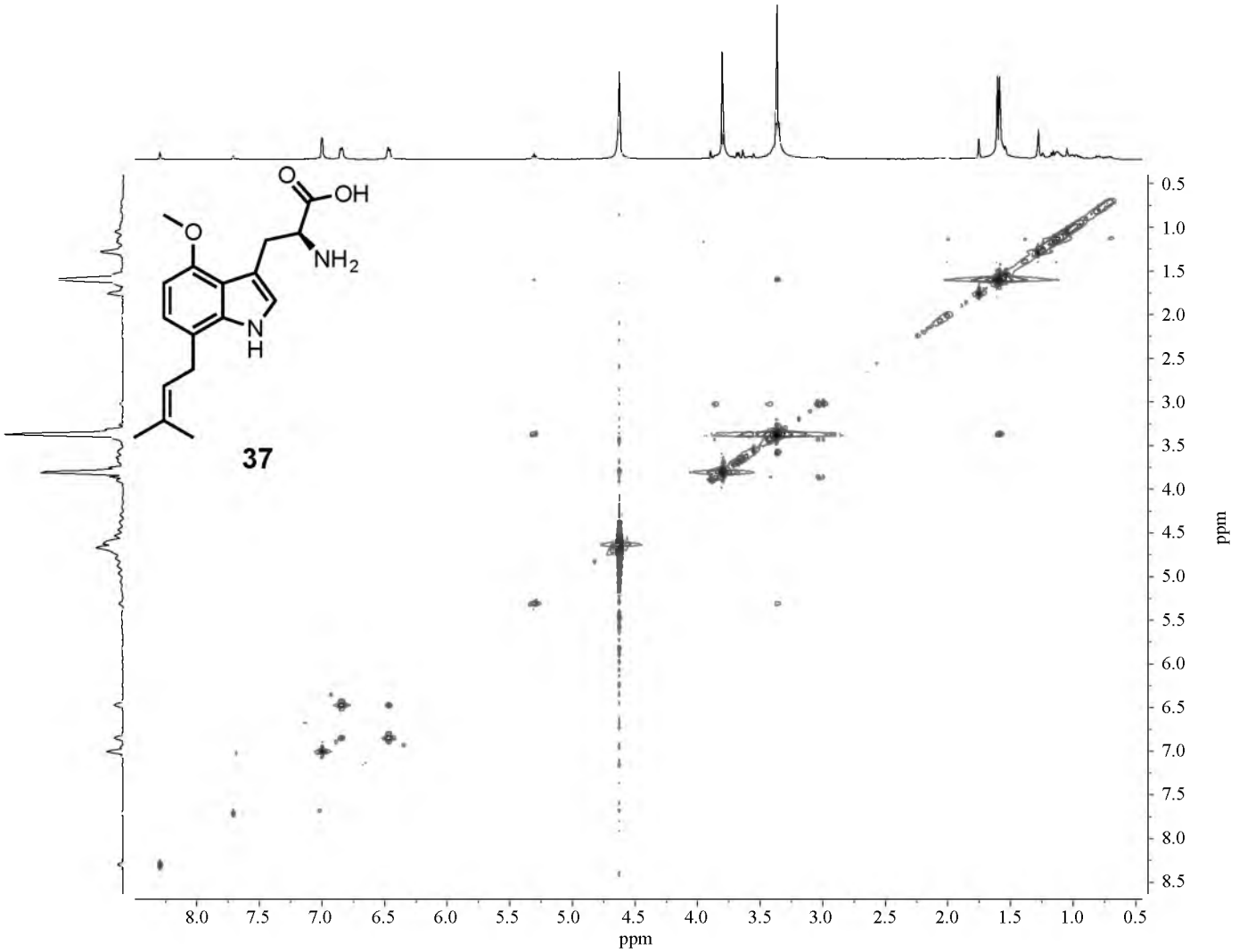
NMR 63. ^{1}H-^{13}C HMBC NMR spectrum of compound 36 in D_2O.



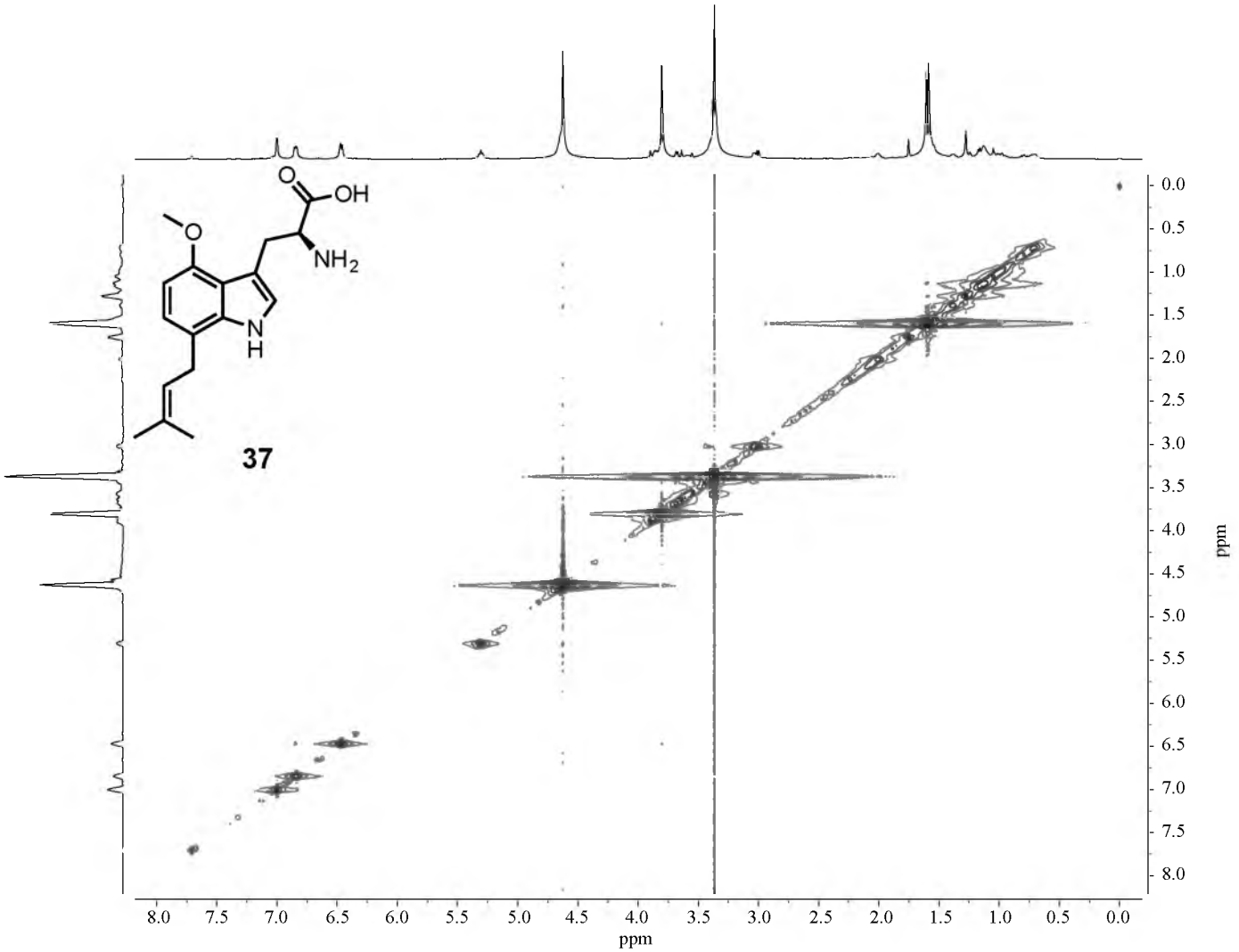
NMR 64. 600 MHz ^{1}H NMR spectrum of compound 37 in D_2O.



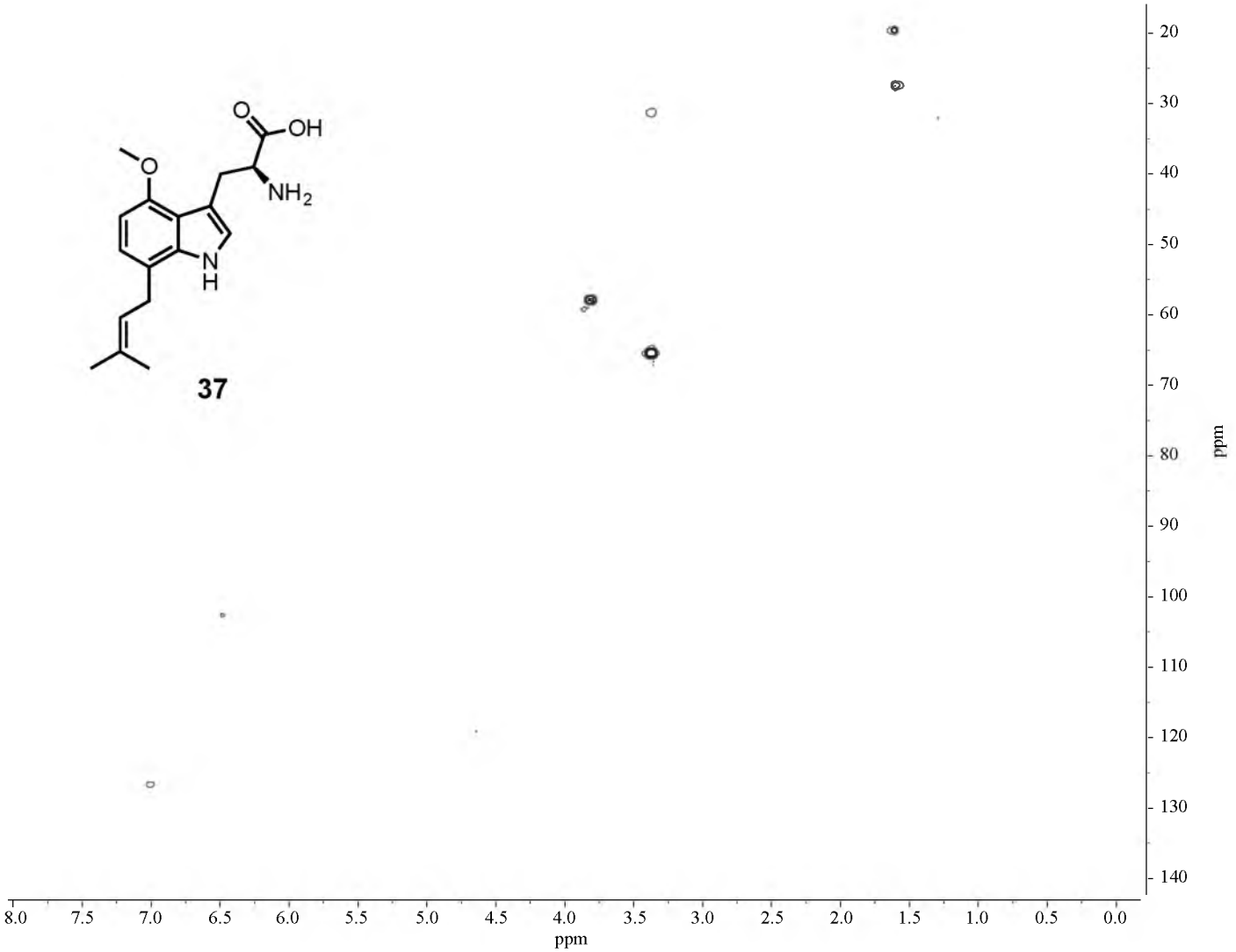
NMR 65. ^{1}H-^{1}H COSY NMR spectrum of compound 37 in D_2O.



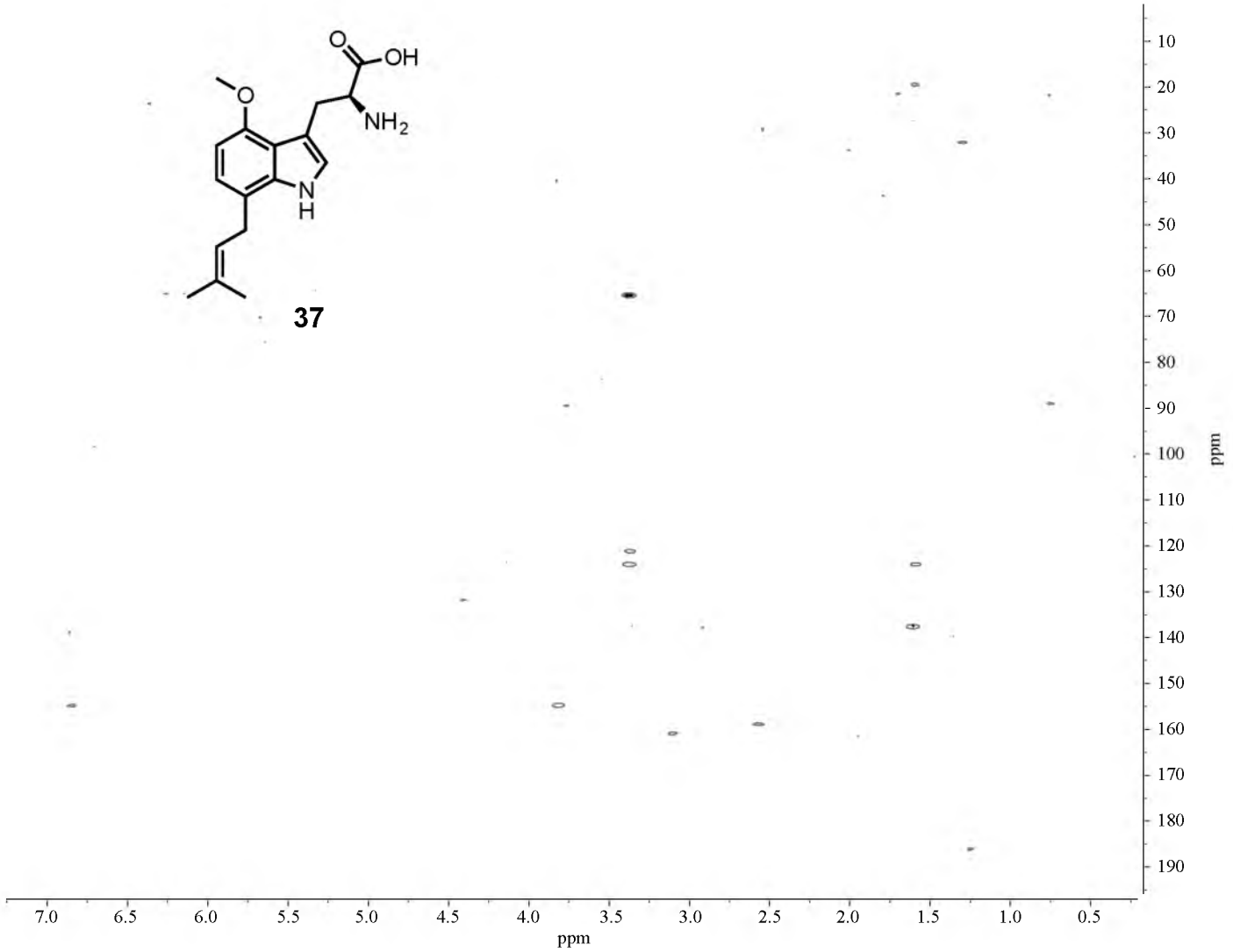
NMR 66. ^{1}H-^{1}H TOCSY NMR spectrum of compound 37 in D_2O.



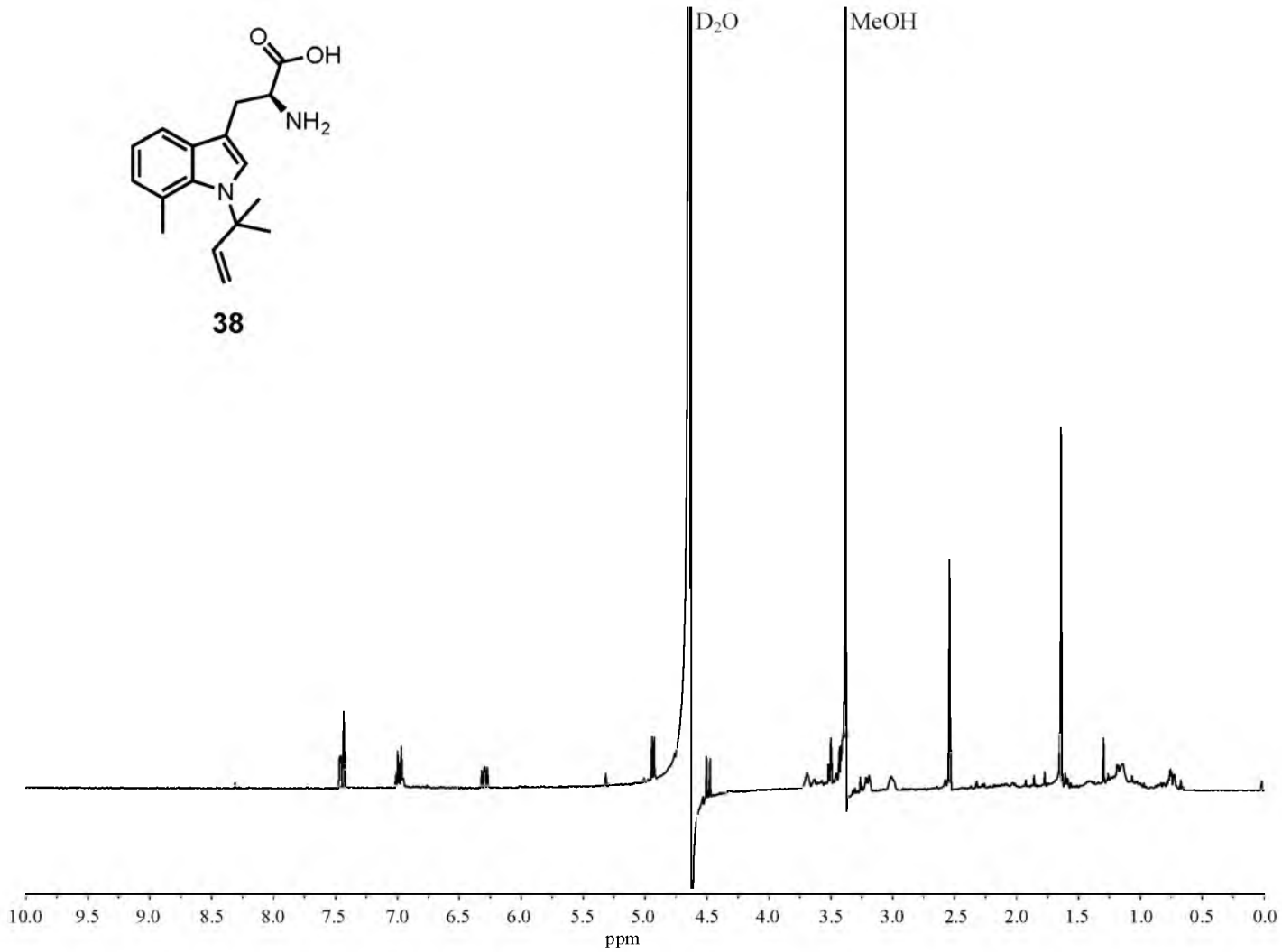
NMR 67. ^{1}H-^{1}H ROESY NMR spectrum of compound 37 in D_2O.



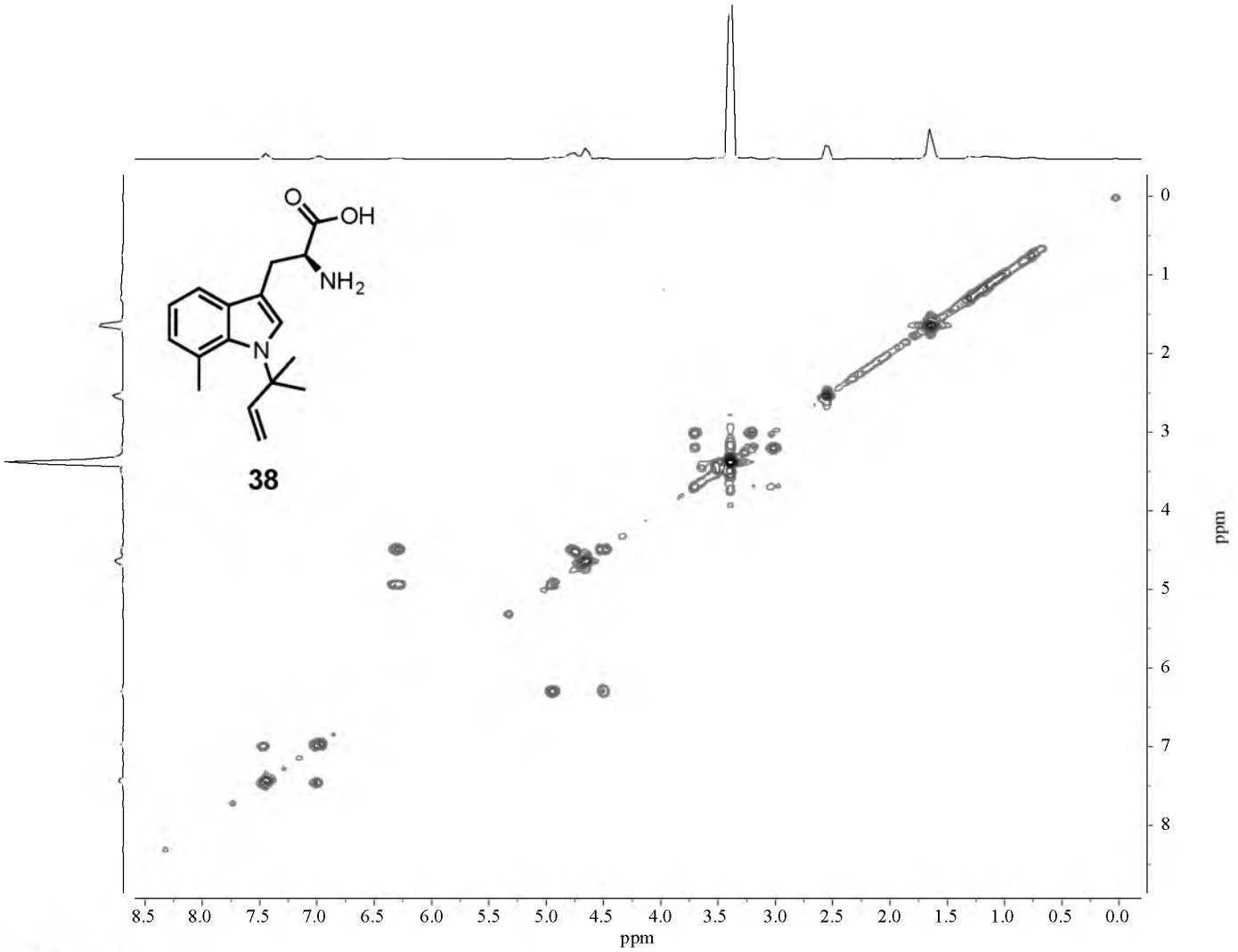
NMR 68. ^{1}H-^{13}C HMQC NMR spectrum of compound 37 in D_2O.



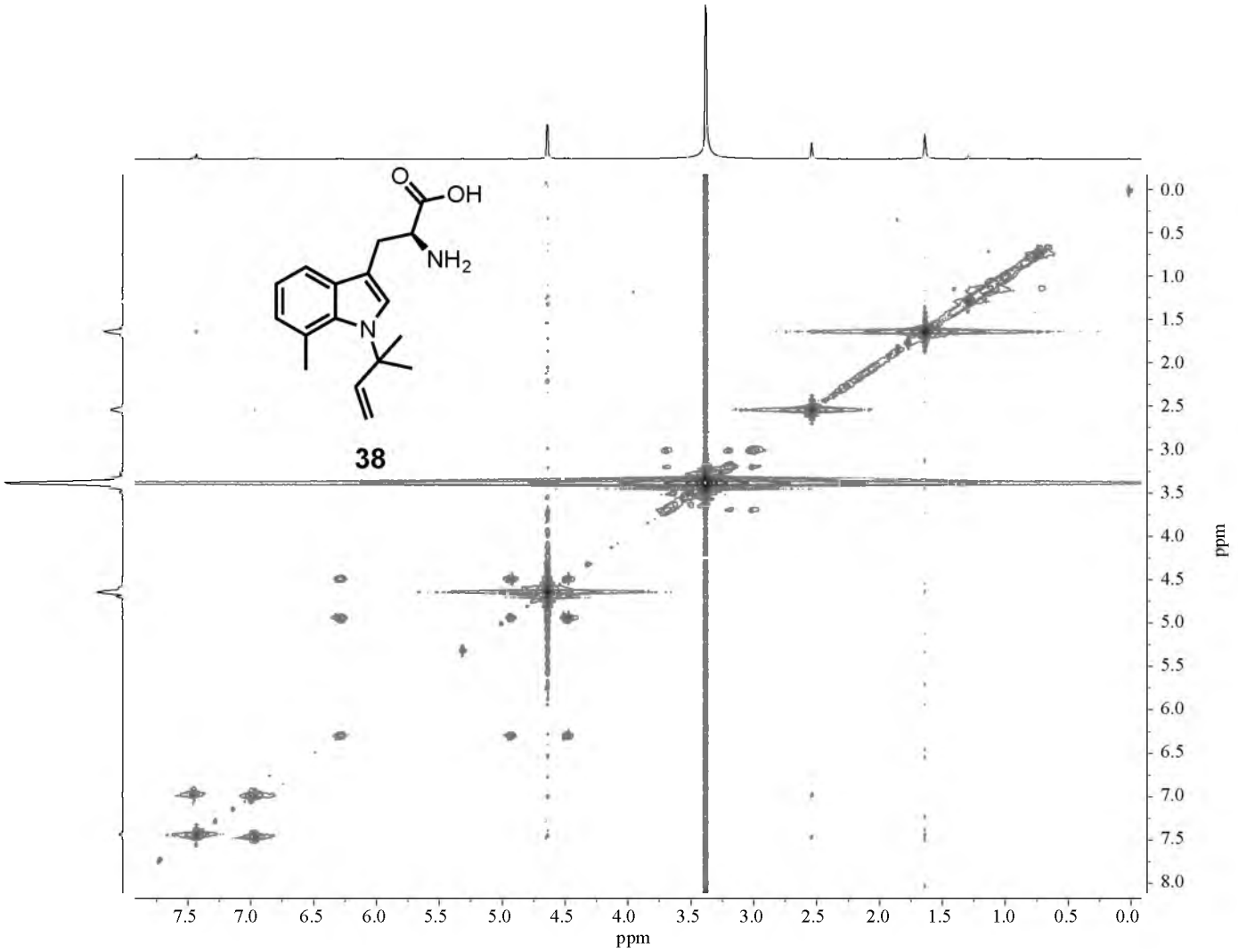
NMR 69. ^{1}H-^{13}C HMBC NMR spectrum of compound 37 in D_2O.



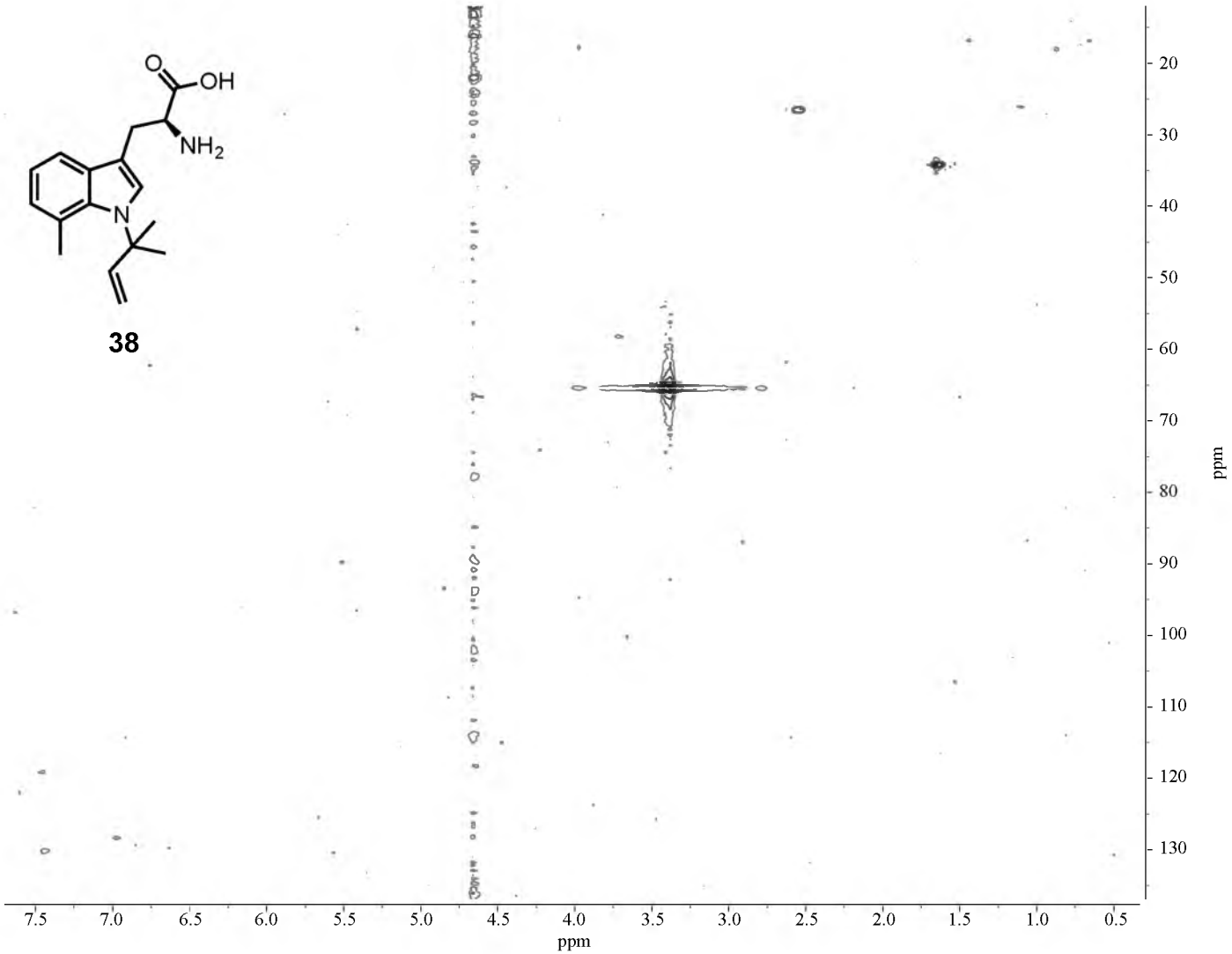
NMR 70. 600 MHz ^{1}H NMR spectrum of compound 38 in D_2O.



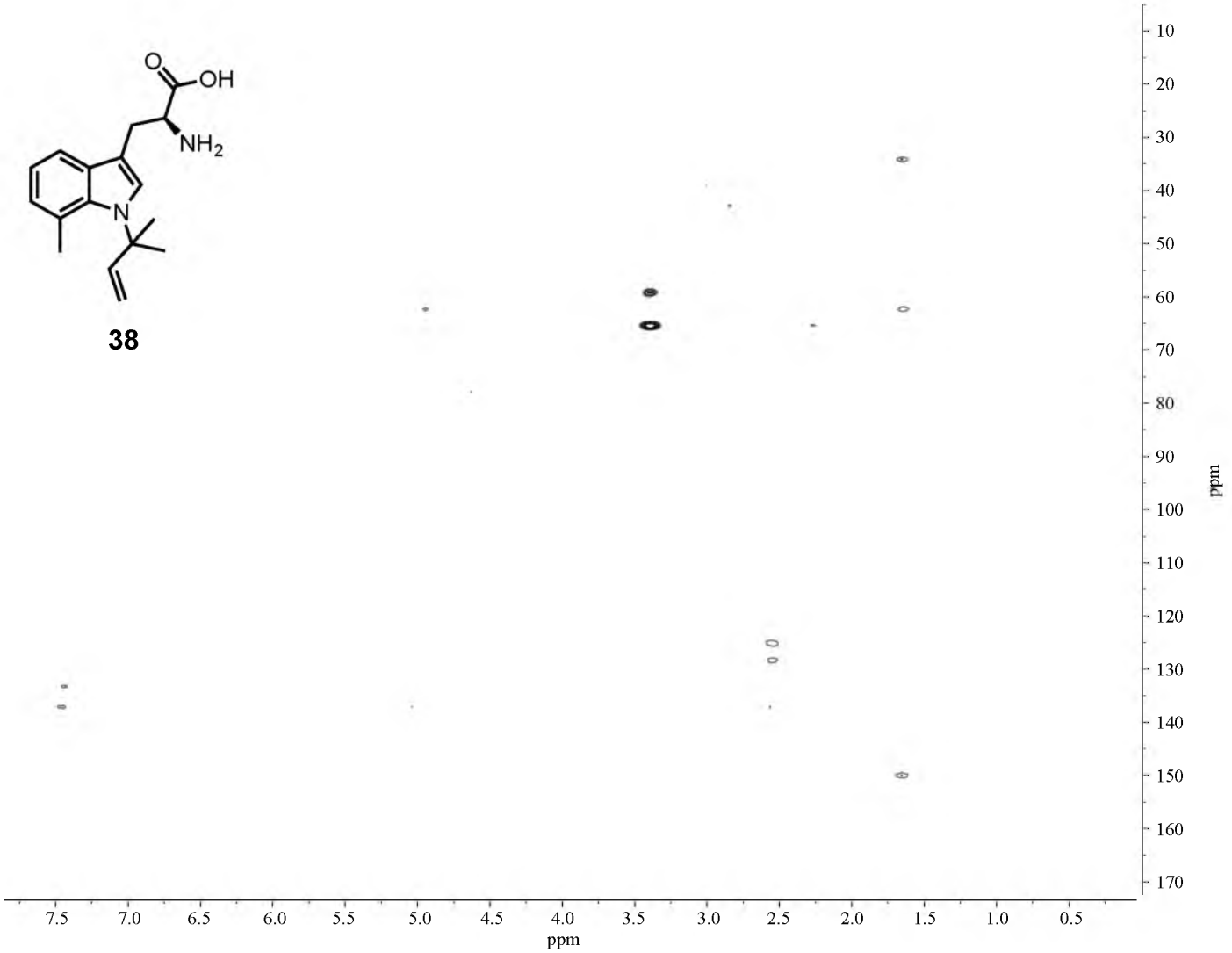
NMR 71. ^{1}H-^{1}H COSY NMR spectrum of compound 38 in D_2O.



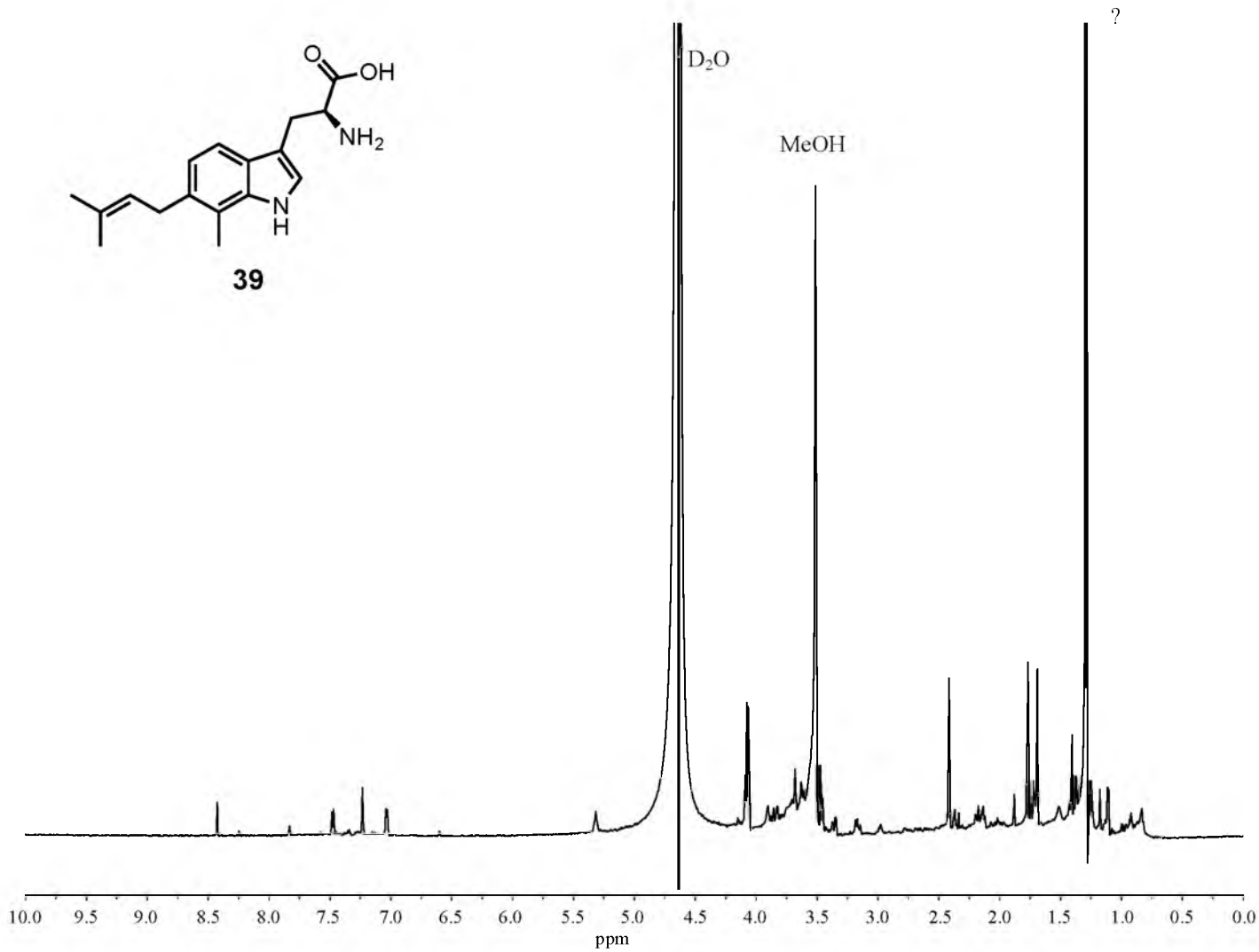
NMR 72. ^{1}H-^{1}H TOCSY NMR spectrum of compound 38 in D_2O.



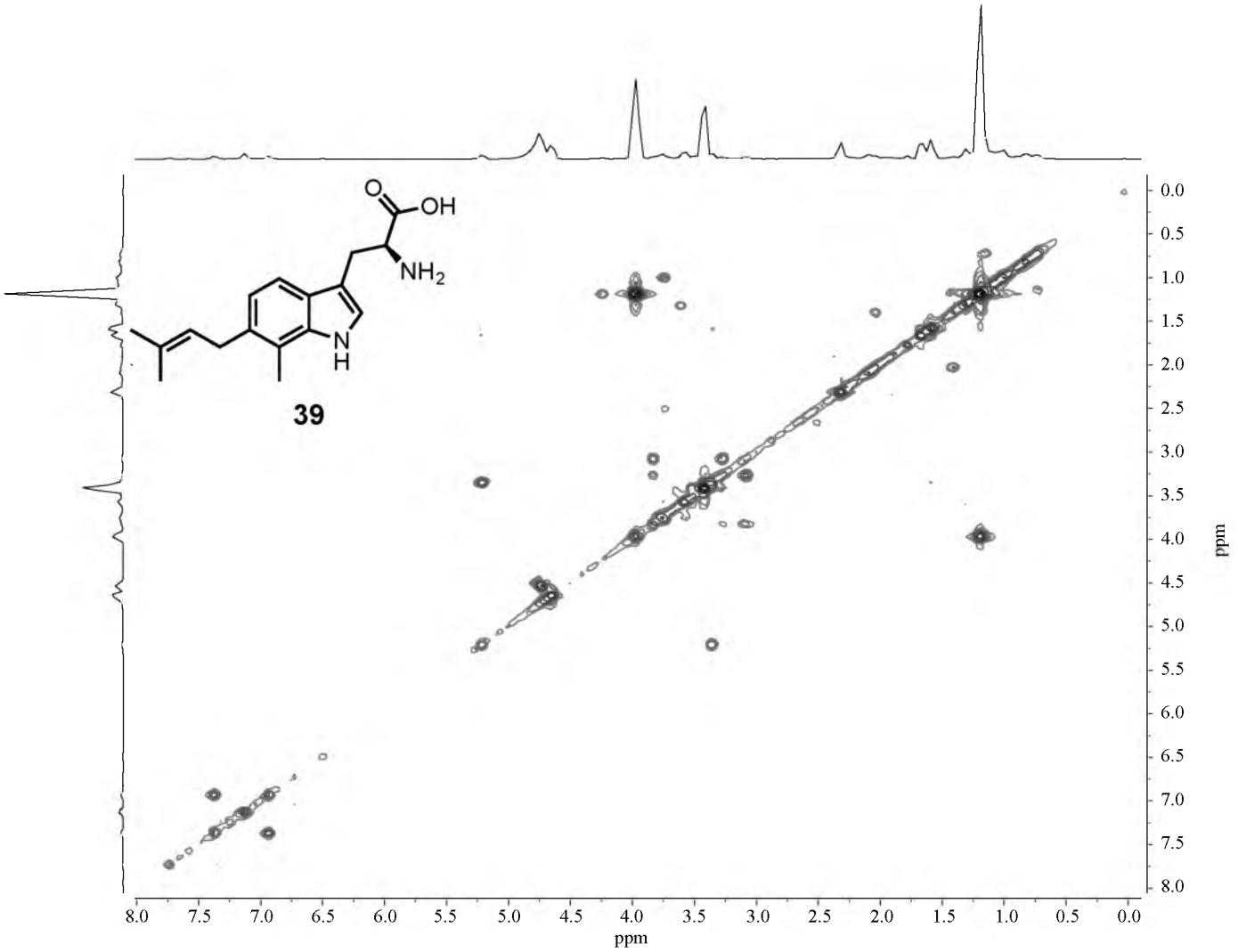
NMR 73. ^{1}H-^{13}C HMQC NMR spectrum of compound 38 in D_2O.



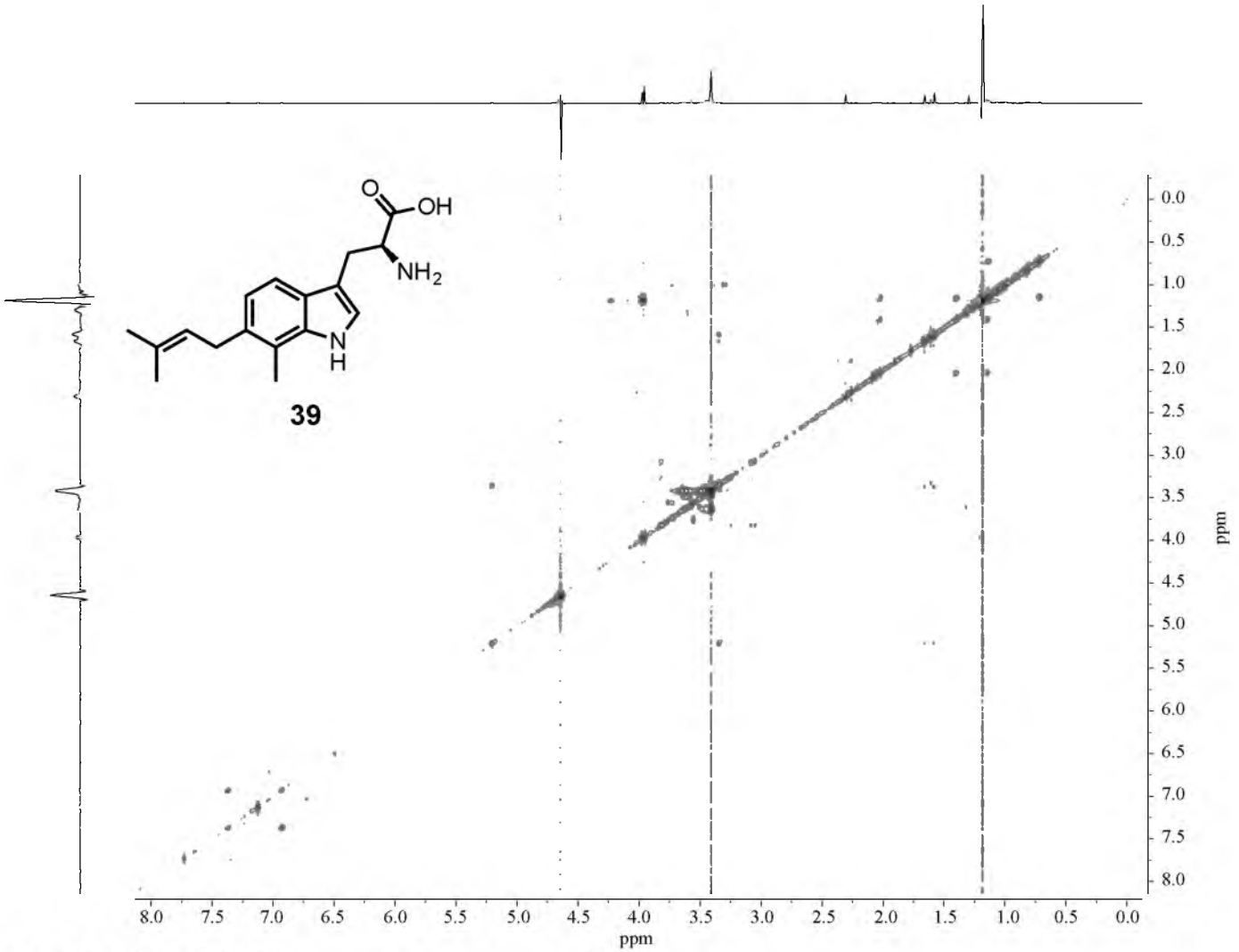
NMR 74. ^{1}H-^{13}C HMBC NMR spectrum of compound 38 in D_2O.



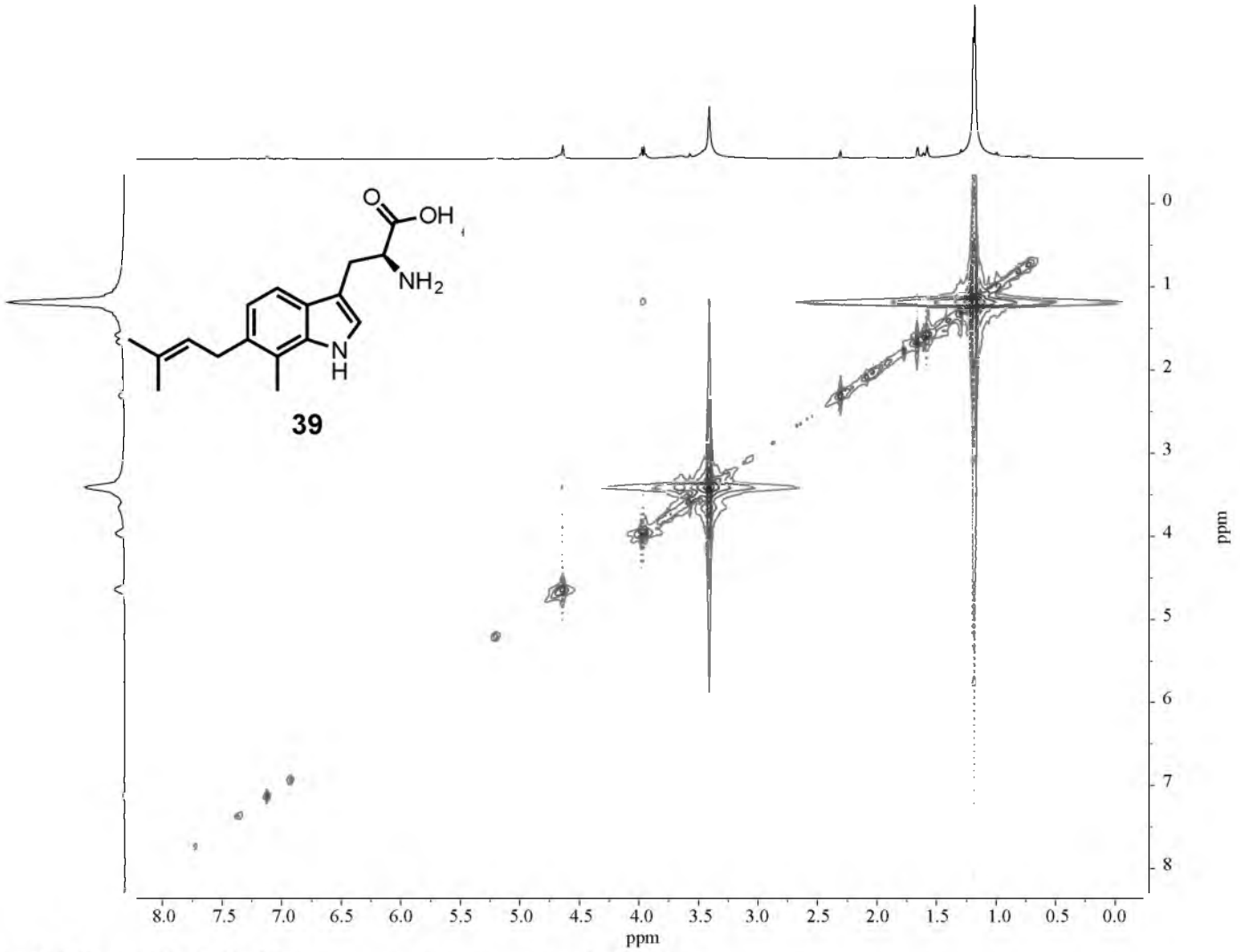
NMR 75. 600 MHz ^{1}H NMR spectrum of compound 39 in D_2O.



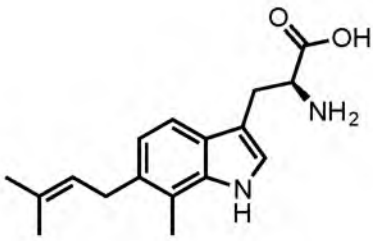
NMR 76. ^{1}H-^{1}H COSY NMR spectrum of compound 39 in D_2O.



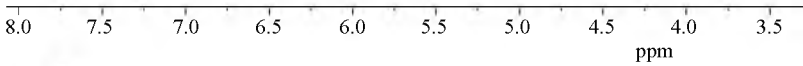
NMR 77. ^{1}H-^{1}H TOCSY NMR spectrum of compound 39 in D_2O.



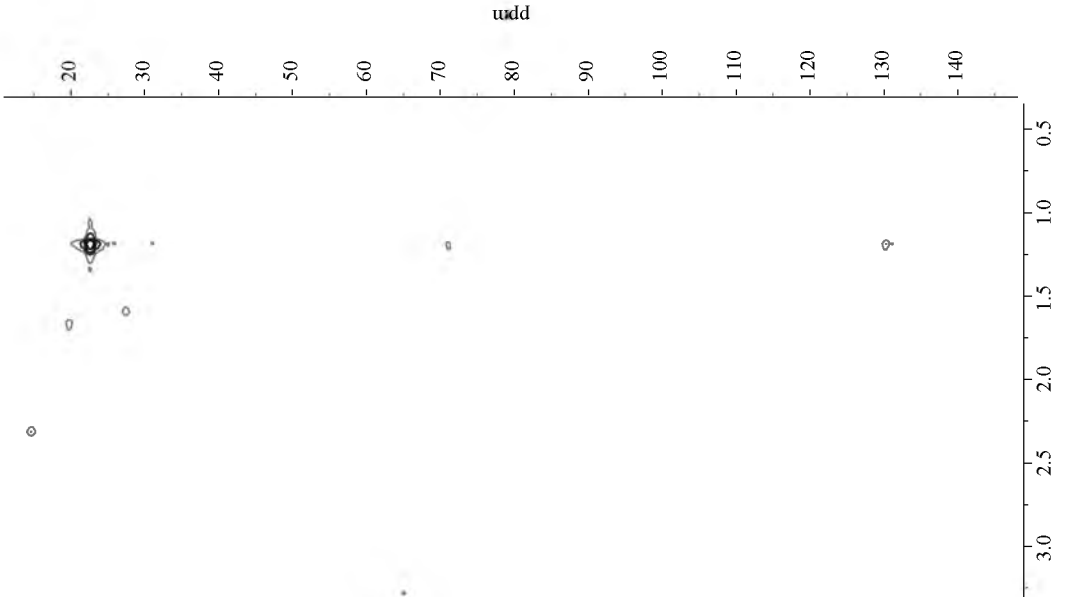
NMR 78. ^{1}H-^{1}H ROESY NMR spectrum of compound 39 in D_2O.

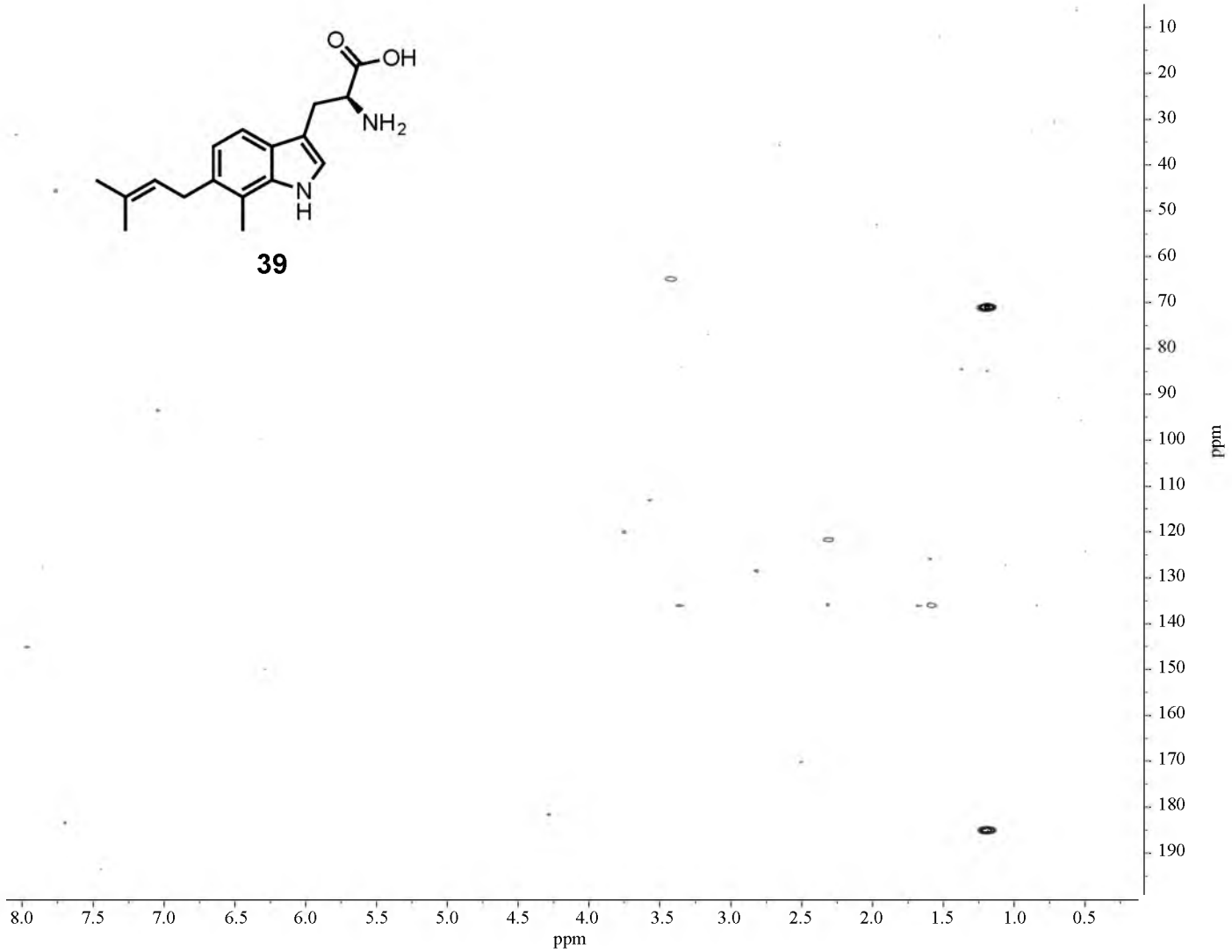


39

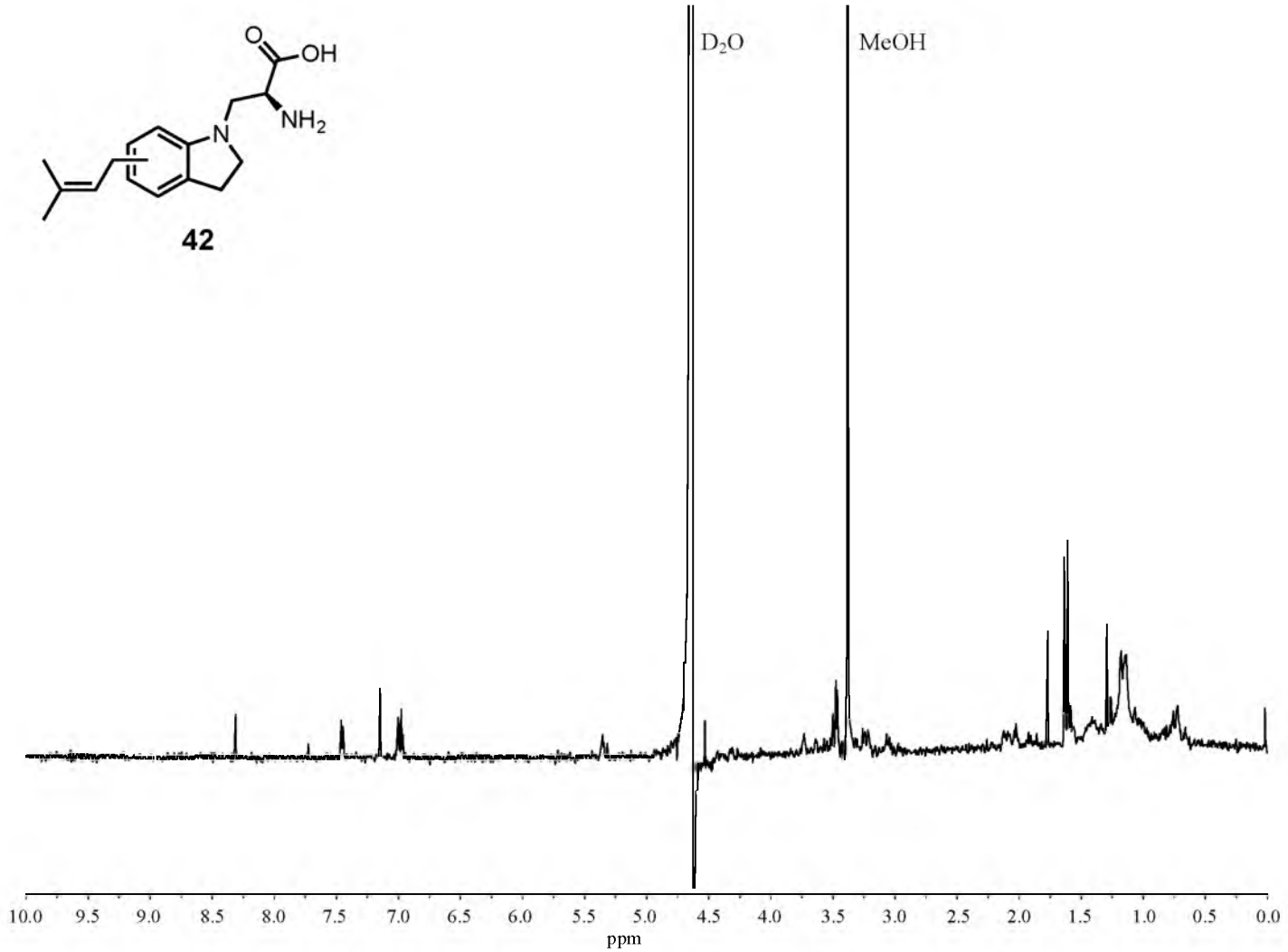


NMR 79. ^{1}H-^{13}C HMQC NMR spectrum of compound 39 in D_2O.

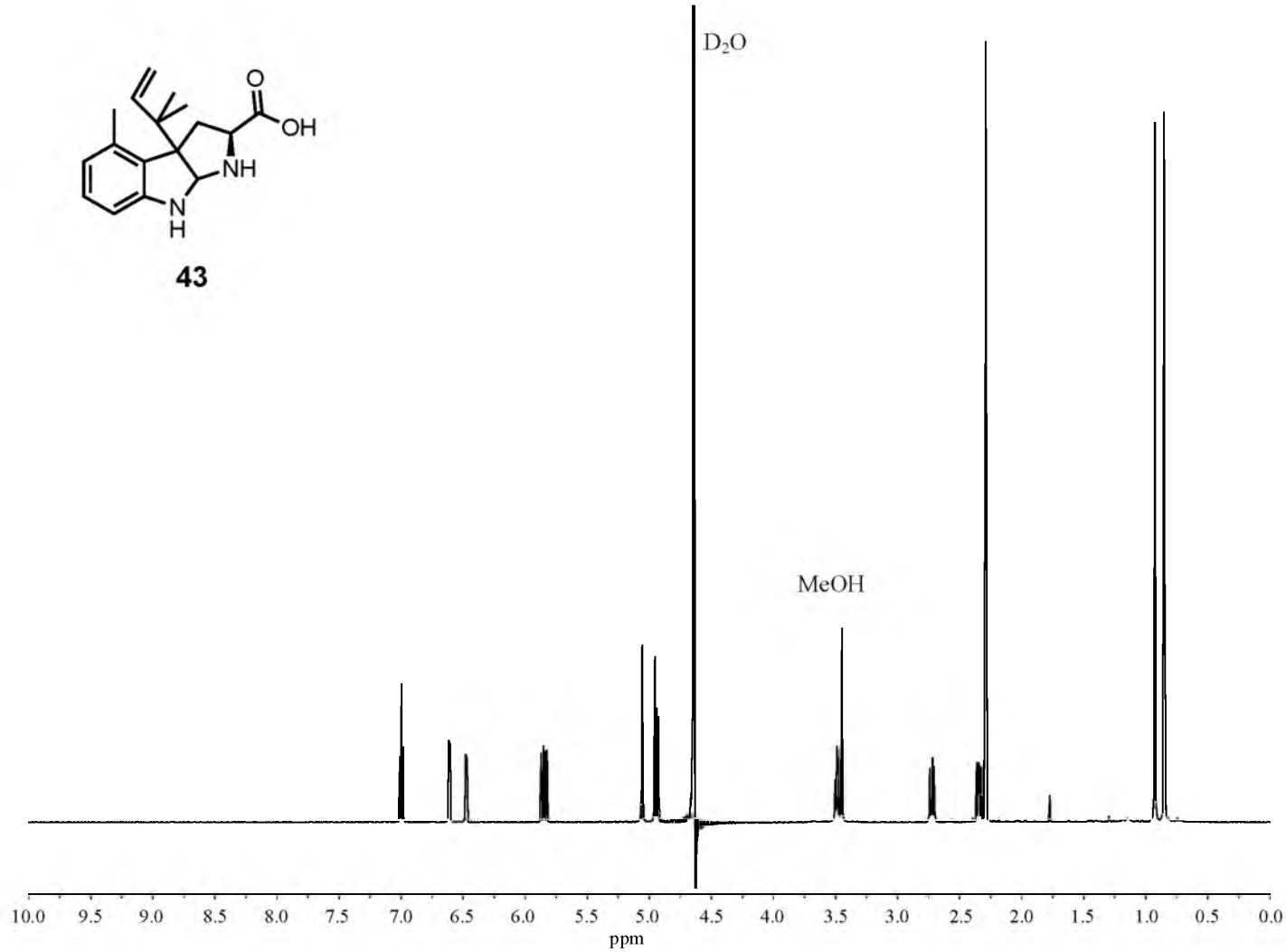

ppm
20
30
40
50
60
70
80
90
100
110
120
130
140
0.5
1.0
1.5
2.0
2.5
3.0



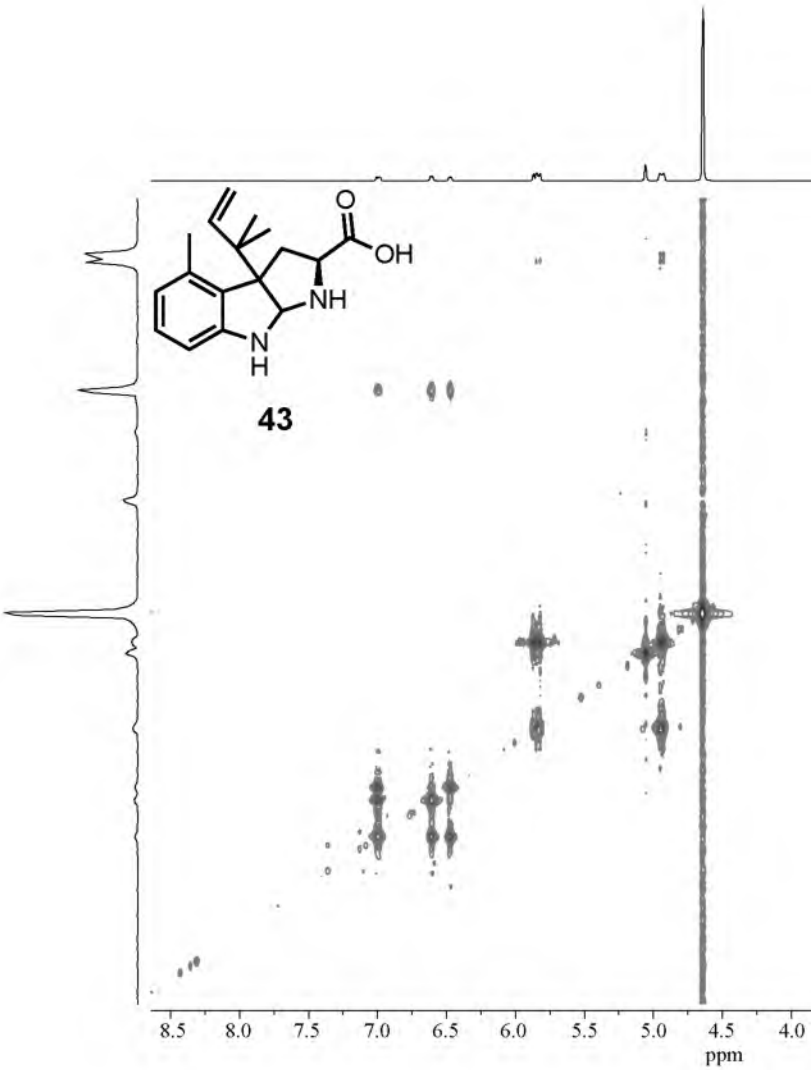
NMR 80. ^{1}H-^{13}C HMBC NMR spectrum of compound 39 in D_2O.



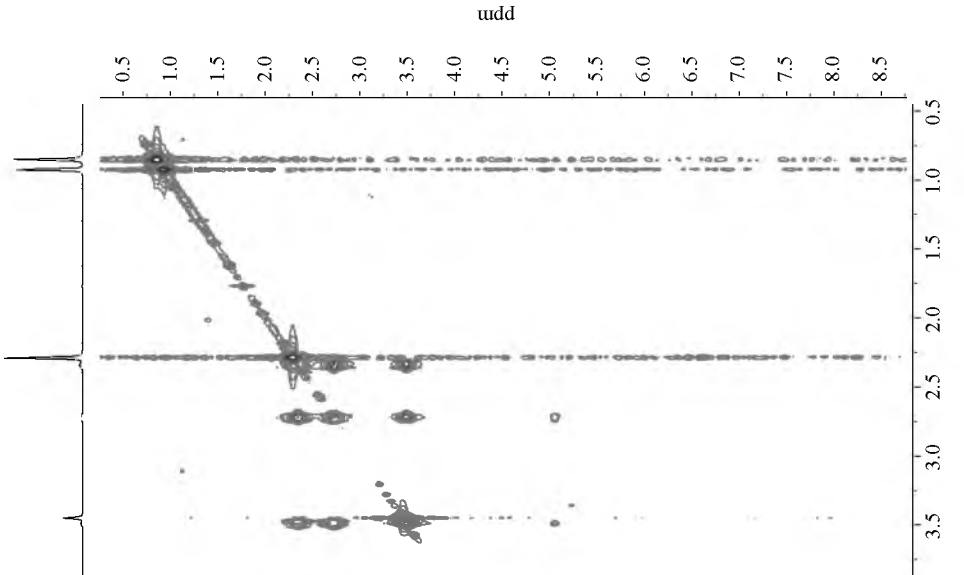
NMR 81. 600 MHz ^{1}H NMR spectrum of compound 42 in D_2O.

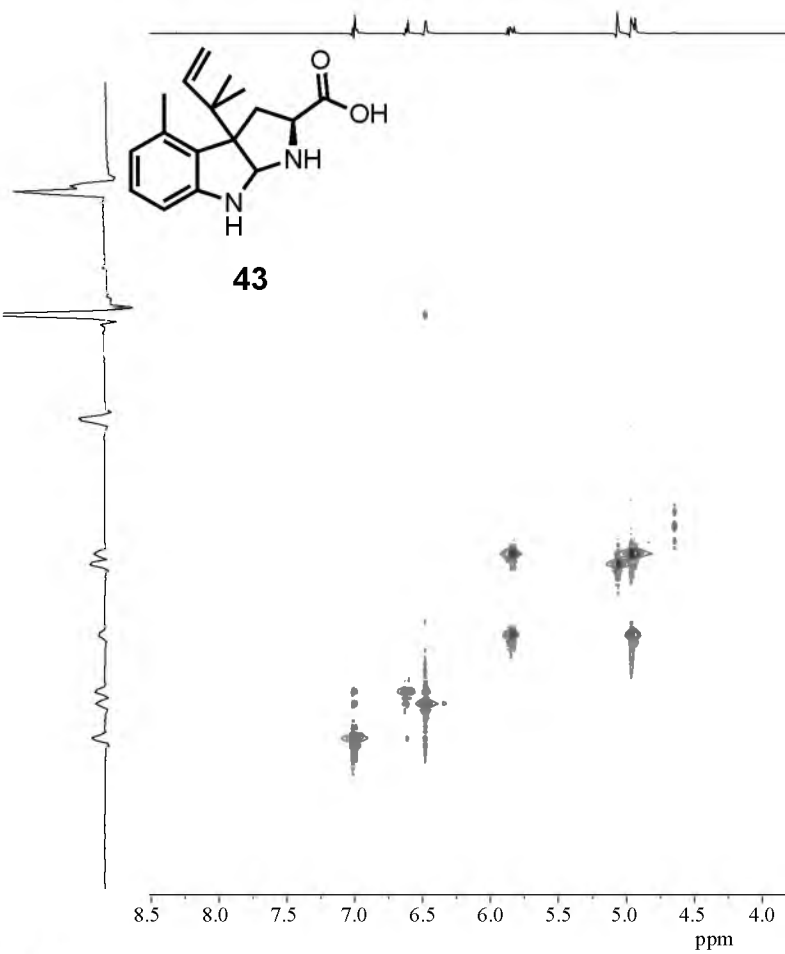


NMR 82. 600 MHz ^{1}H NMR spectrum of compound 43 in D_2O.



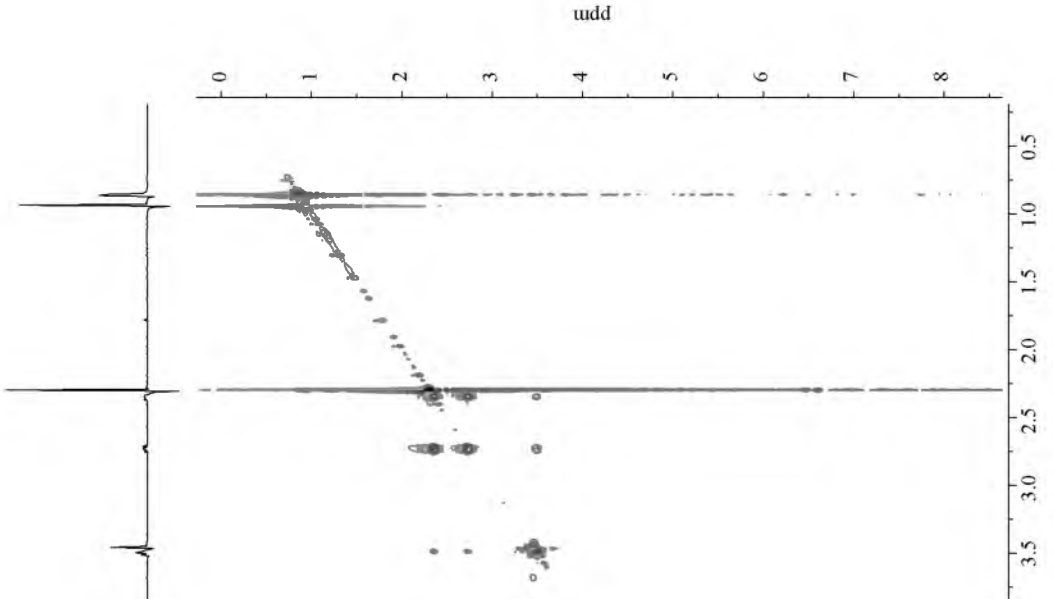
NMR 83. ^{1}H-^{1}H COSY NMR spectrum of compound 43 in D_2O.

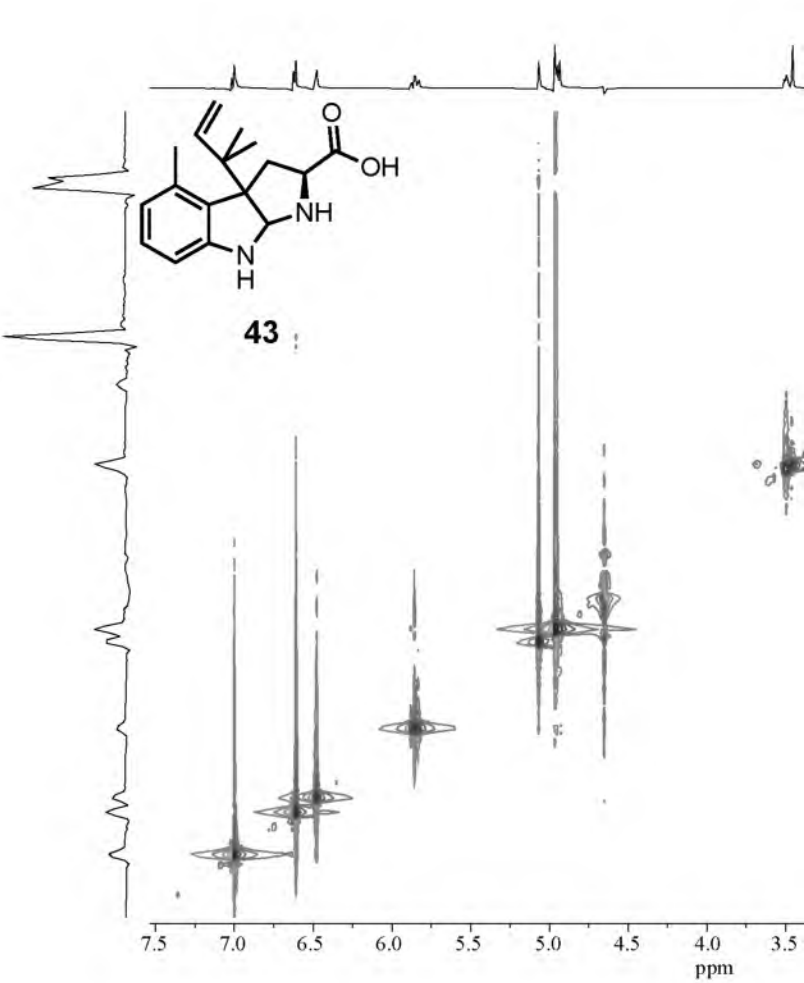

ppm
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5



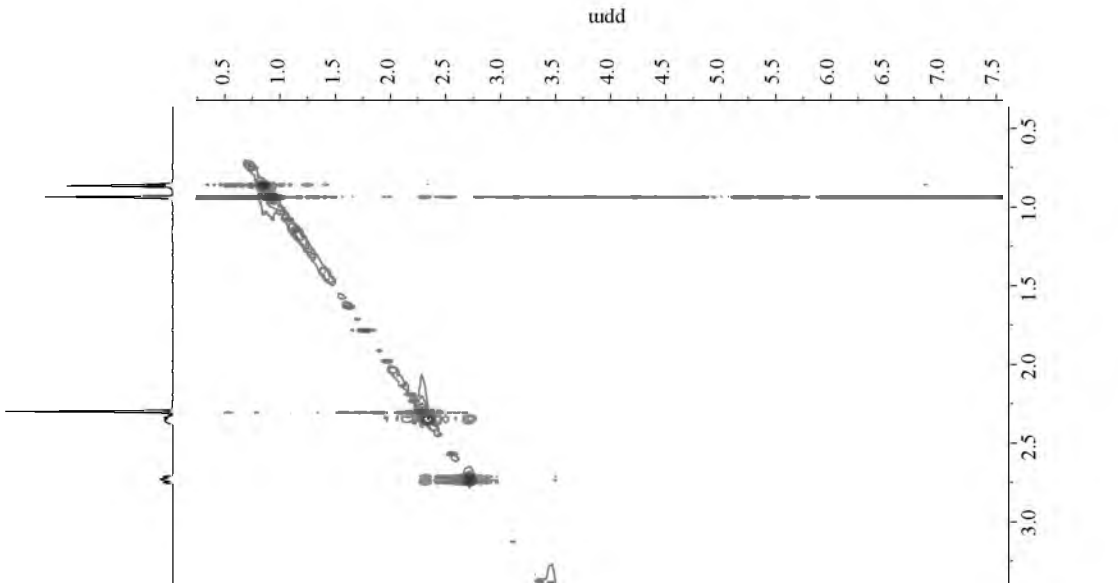
43

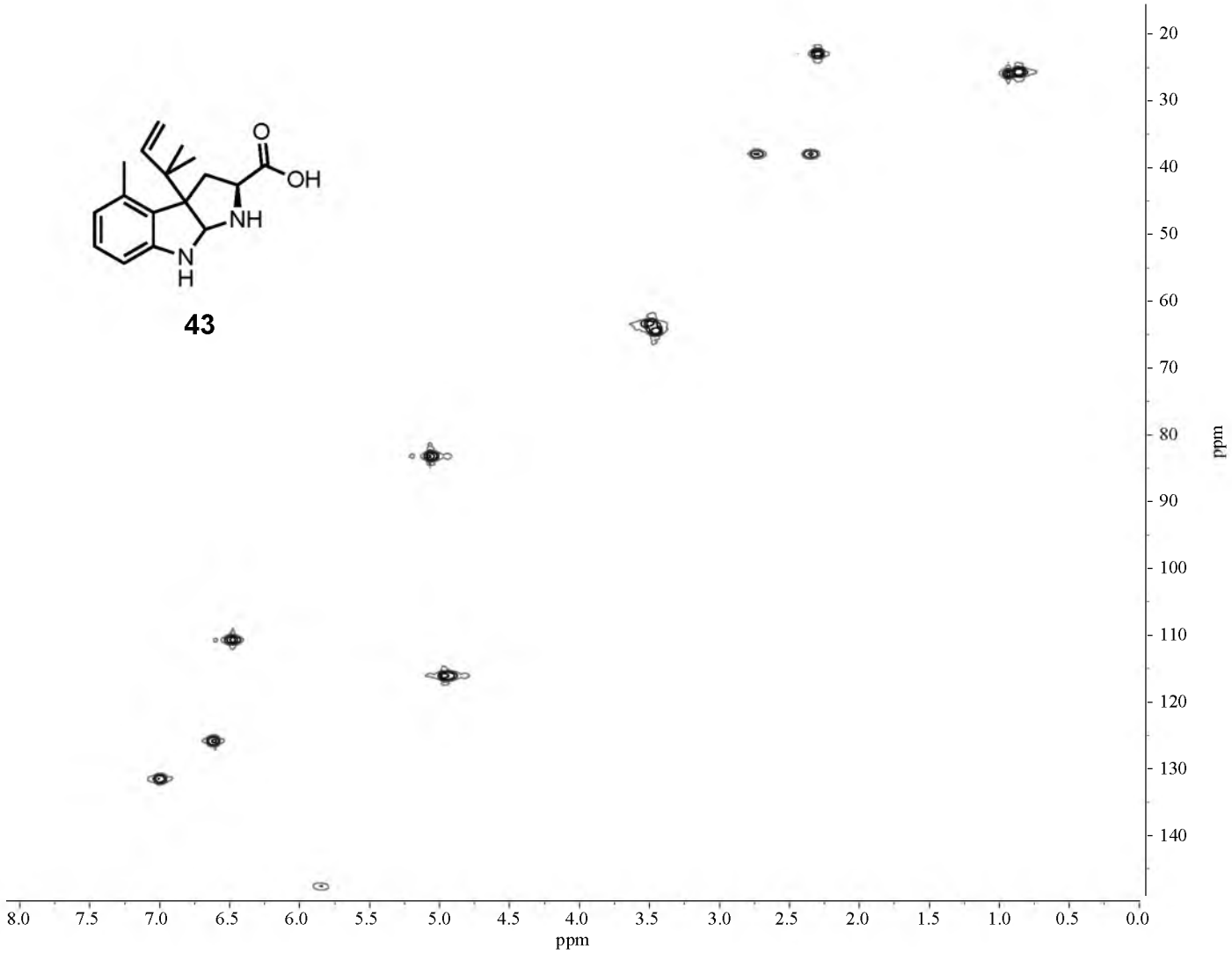
NMR 84. ^{1}H-^{1}H TOCSY NMR spectrum of compound 43 in D_2O.


ppm
0
1
2
3
4
5
6
7
8
0.5
1.0
1.5
2.0
2.5
3.0
3.5

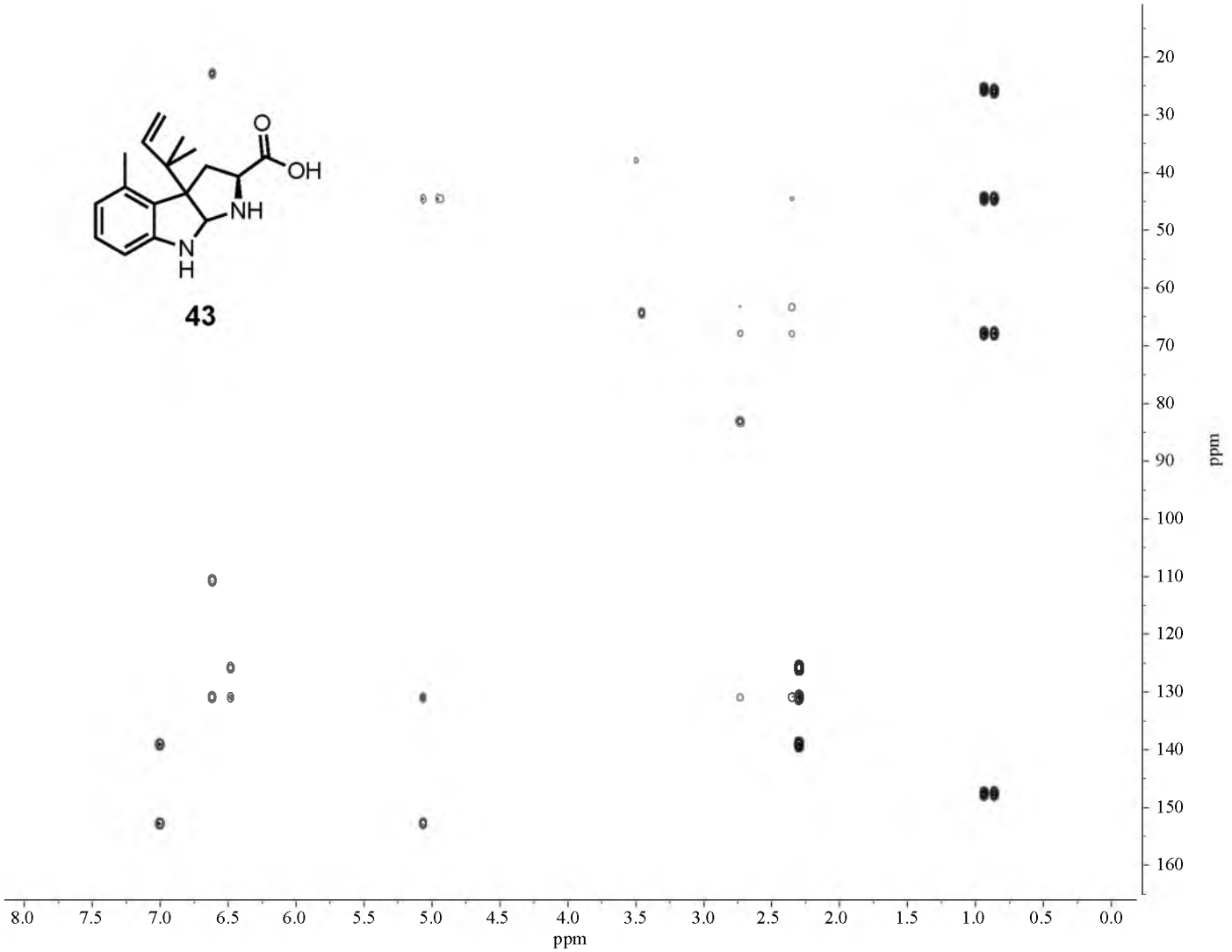


NMR 85. ^{1}H-^{1}H ROESY NMR spectrum of compound 43 in D_2O.

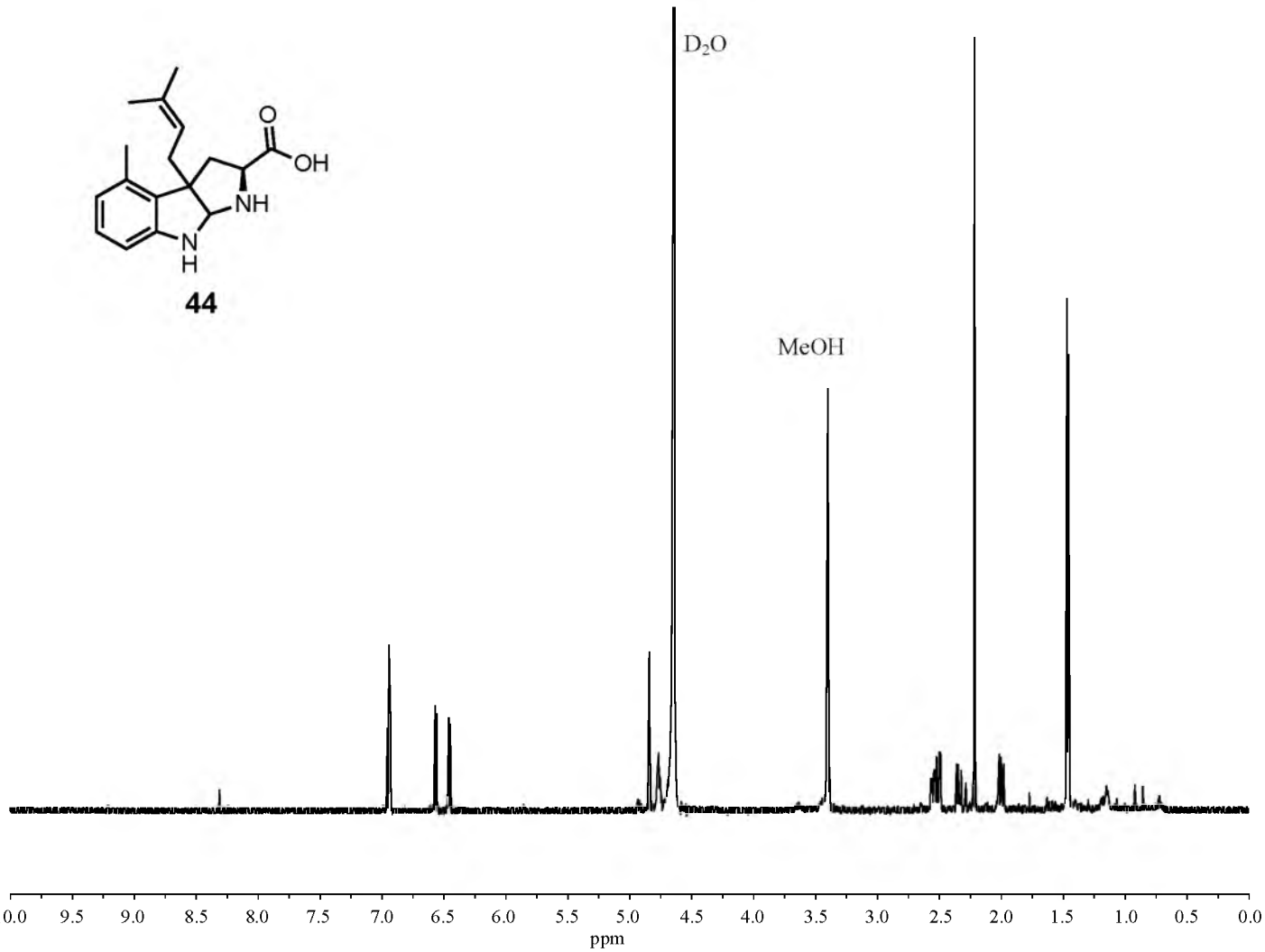

ppm
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5



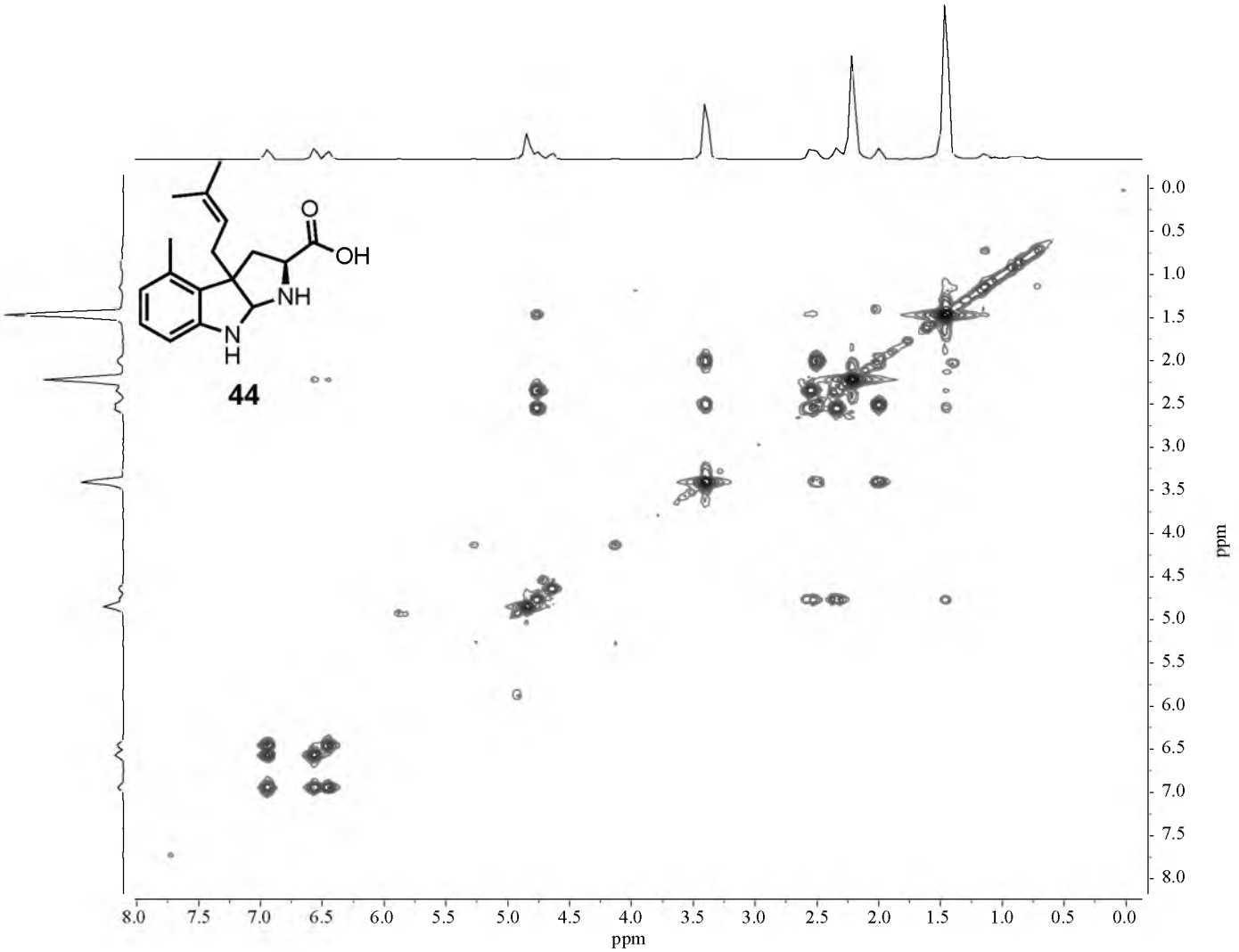
NMR 86. ^{1}H-^{13}C HSQC NMR spectrum of compound 43 in D_2O.



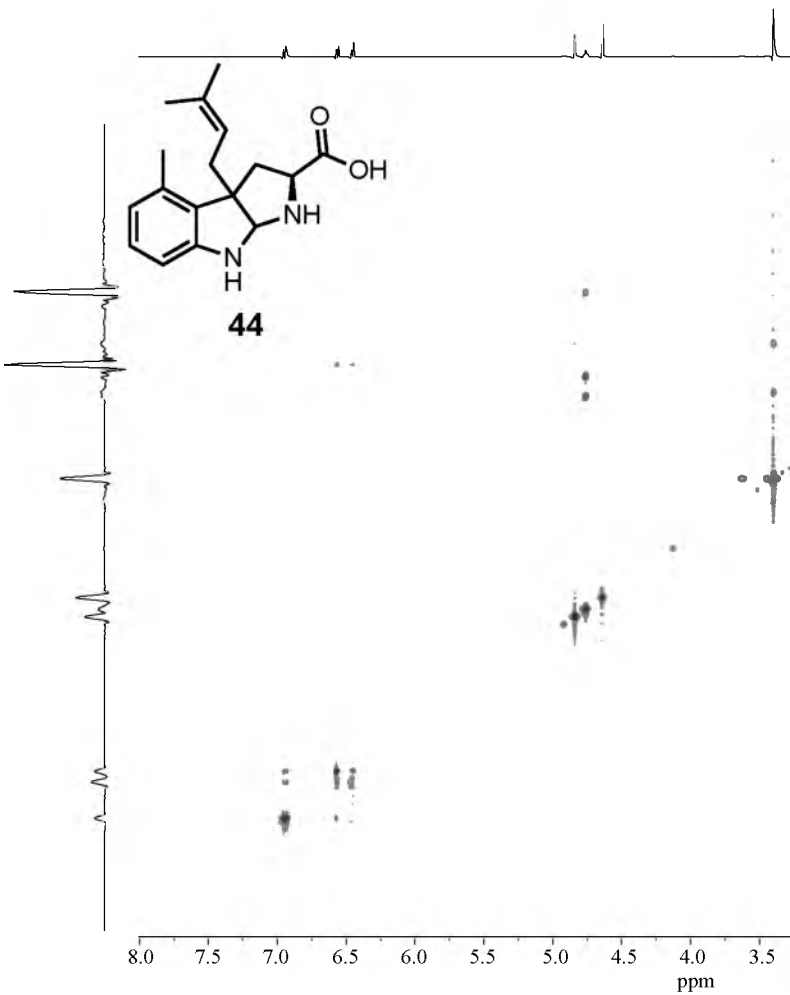
NMR 87. ^{1}H-^{13}C HMBC NMR spectrum of compound 43 in D_2O.



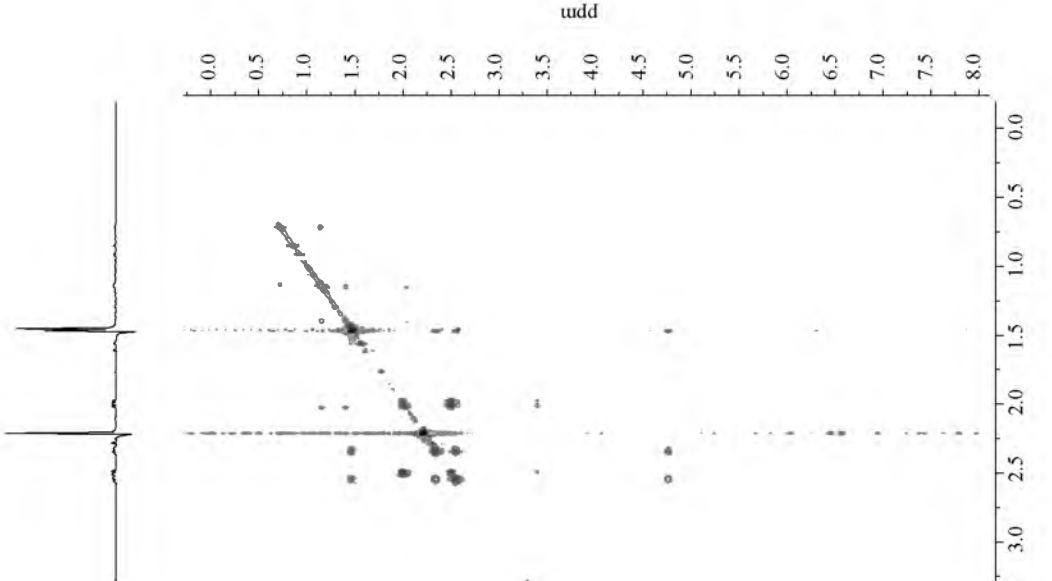
NMR 88. 600 MHz ^{1}H NMR spectrum of compound 44 in D_2O.

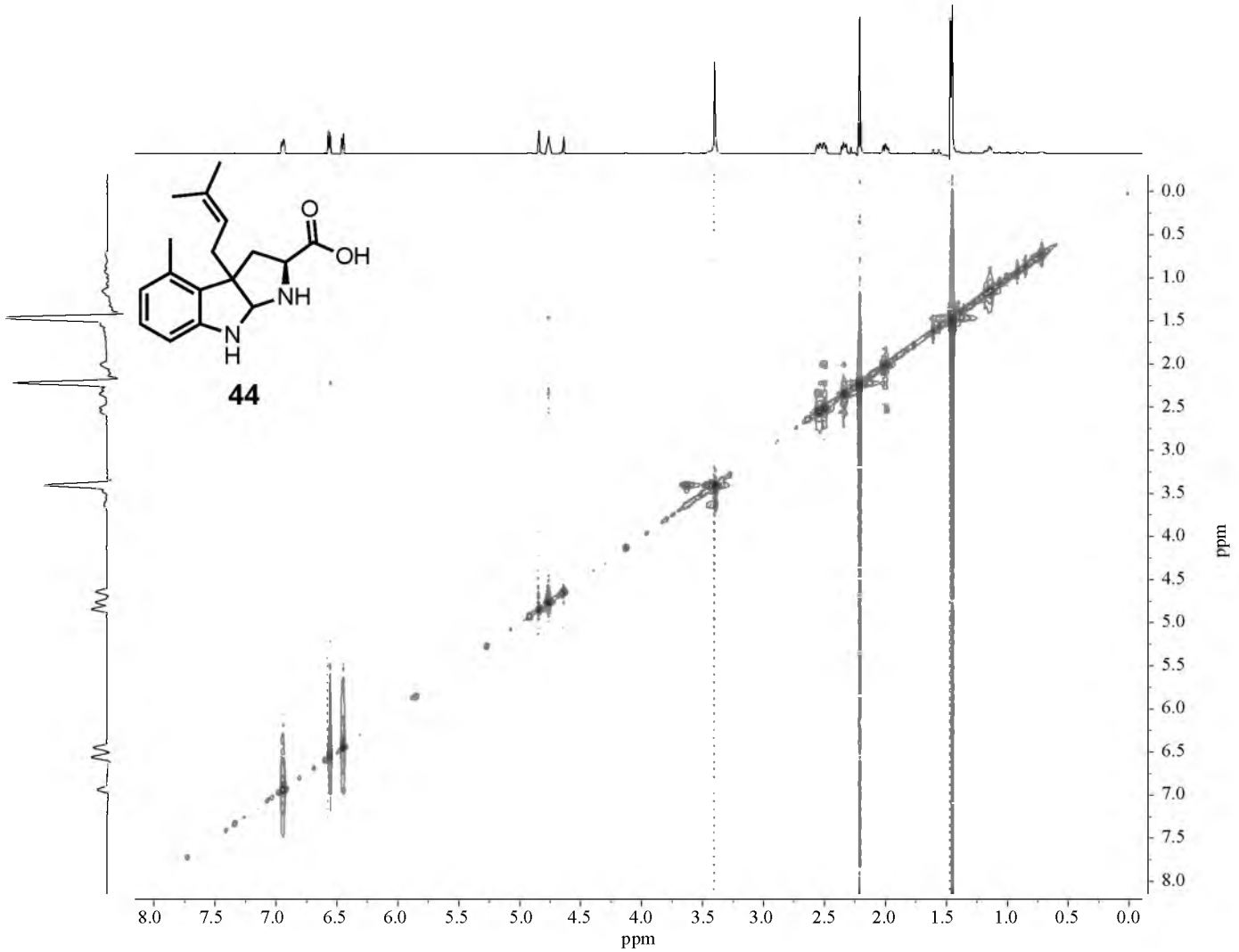


NMR 89. ^{1}H-^{1}H COSY NMR spectrum of compound 44 in D_2O.

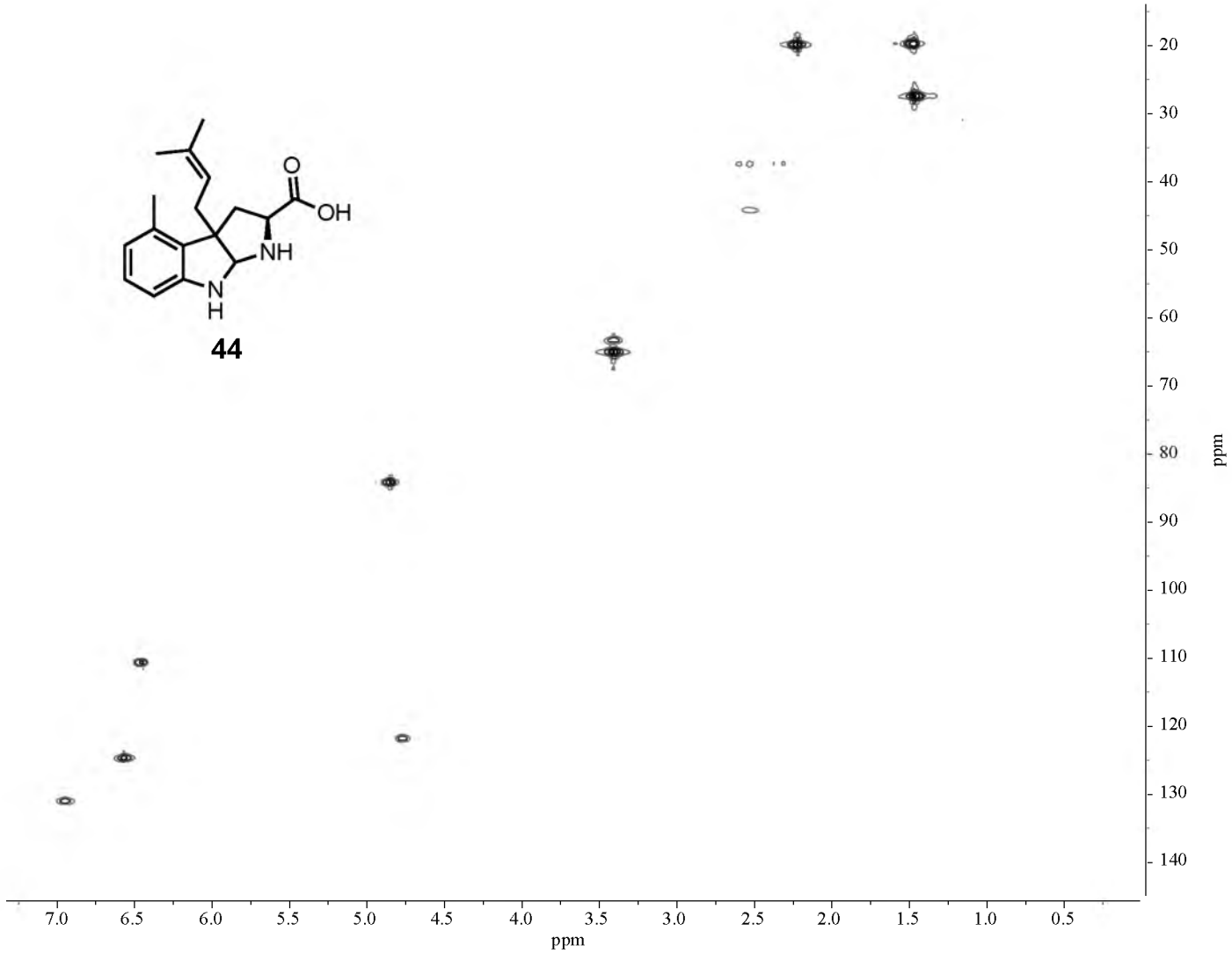


NMR 90. ^{1}H-^{1}H TOCSY NMR spectrum of compound 44 in D_2O.

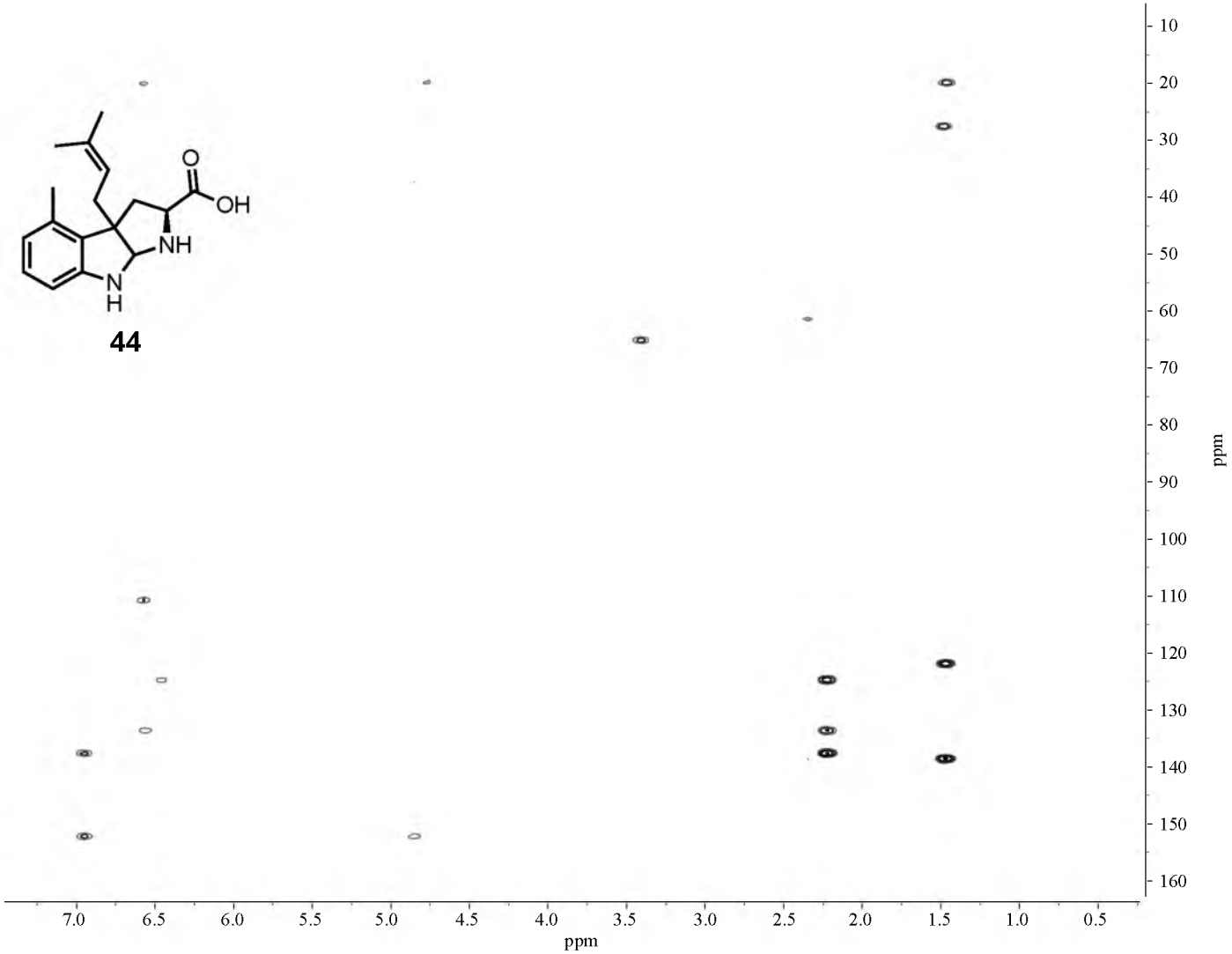

ppm
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0



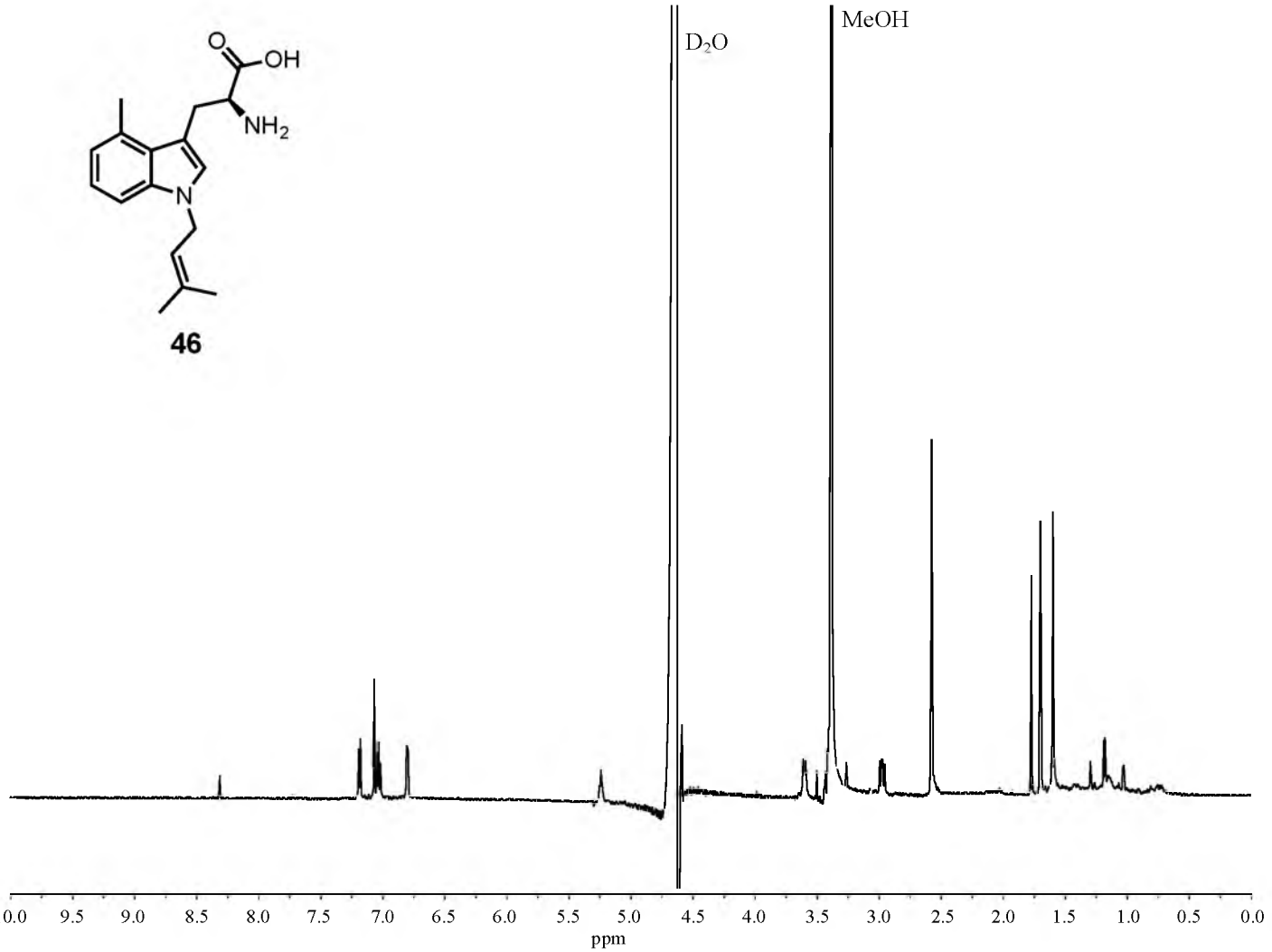
NMR 91. ^{1}H-^{1}H ROESY NMR spectrum of compound 44 in D_2O.



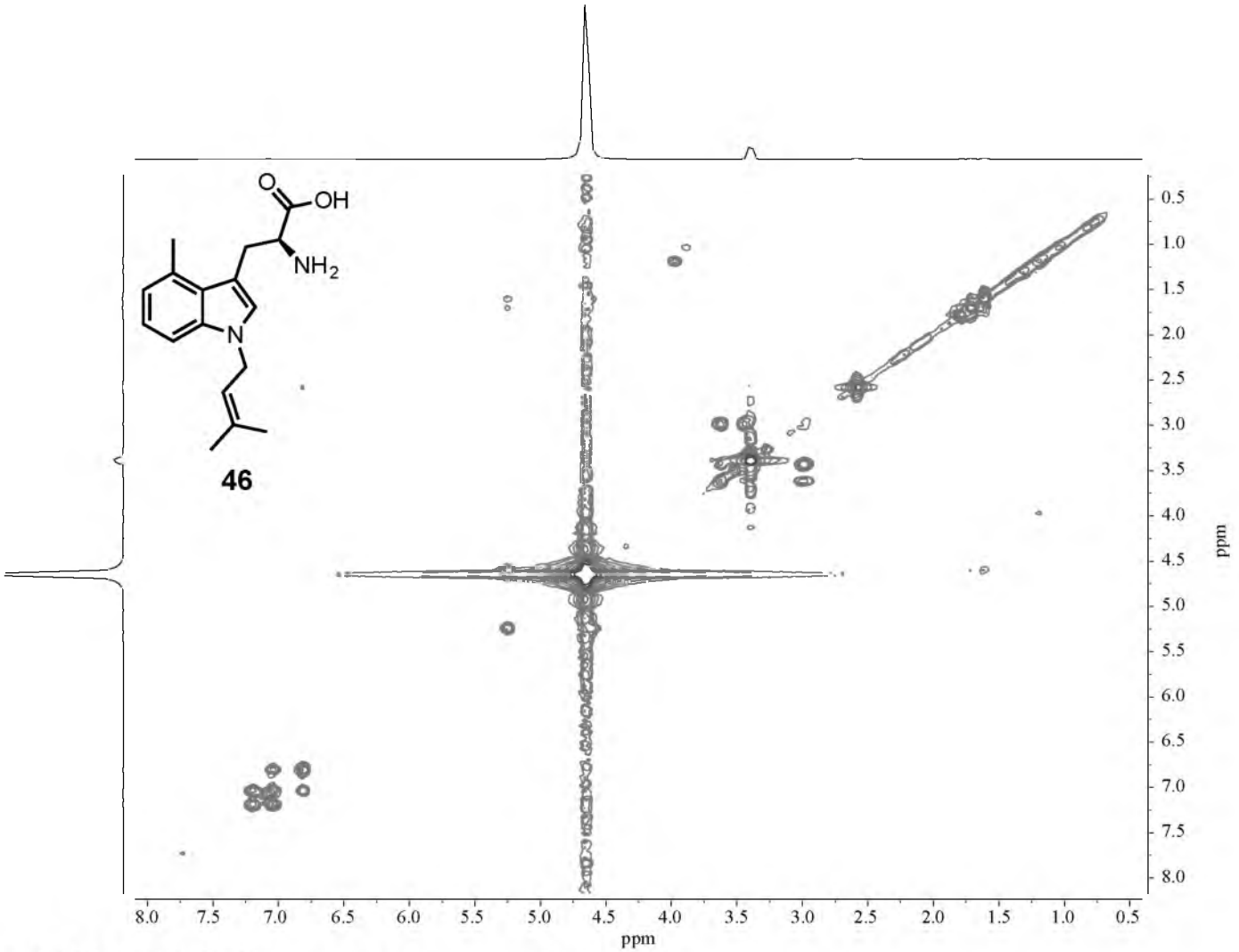
NMR 92. ^{1}H-^{13}C HMQC NMR spectrum of compound 44 in D_2O.



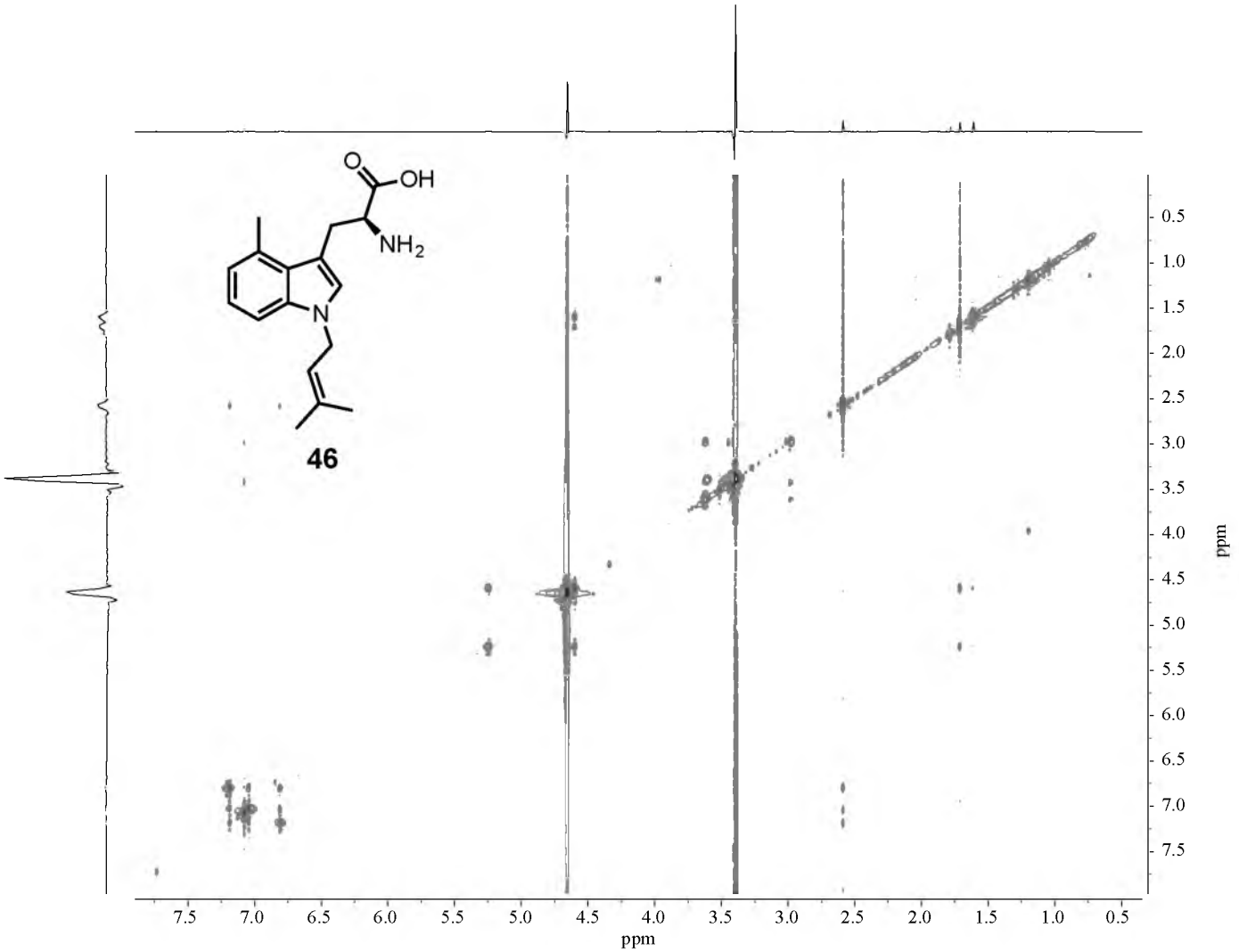
NMR 93. ^{1}H-^{13}C HMBC NMR spectrum of compound 44 in D_2O.



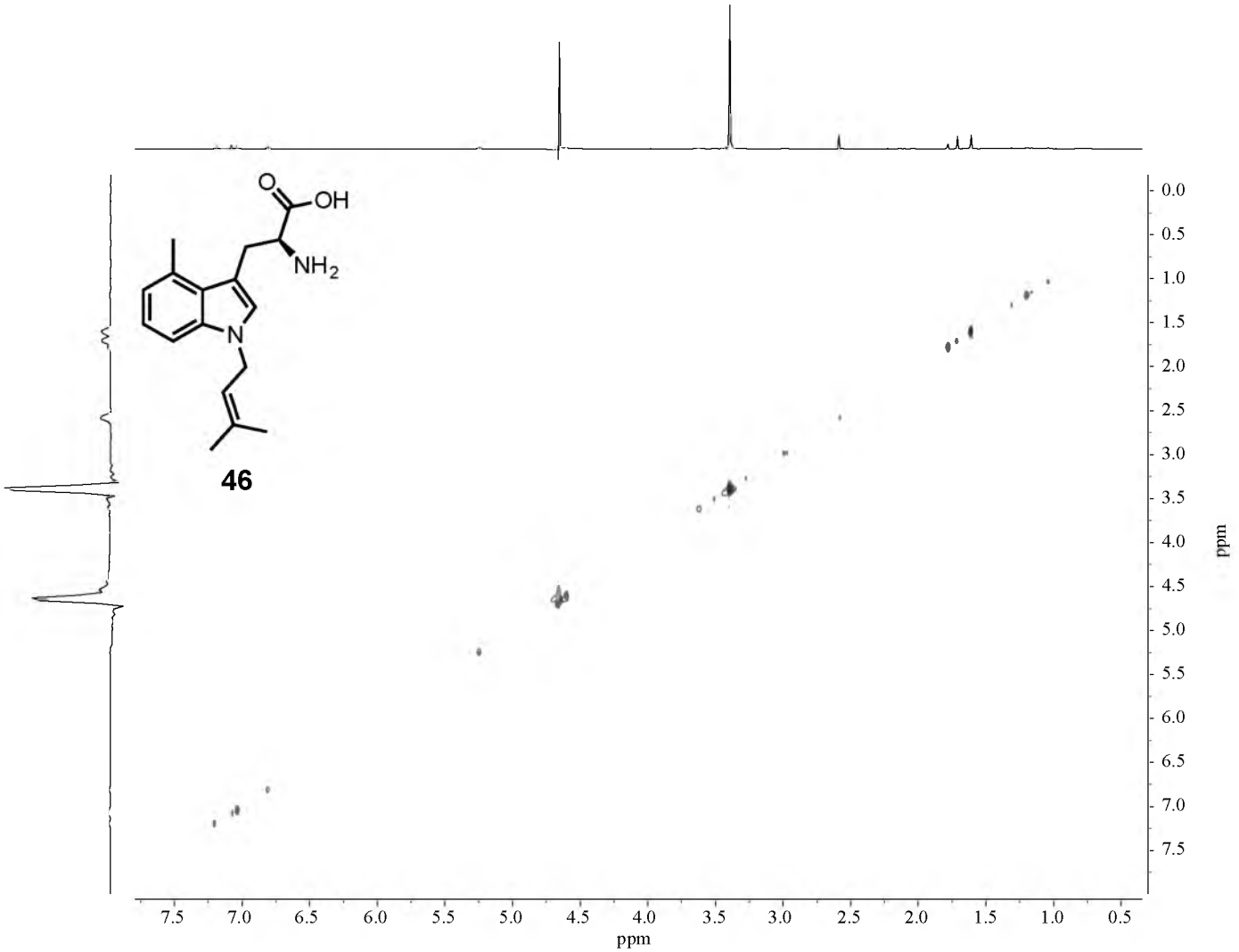
NMR 94. 600 MHz ^{1}H NMR spectrum of compound 46 in D_2O.



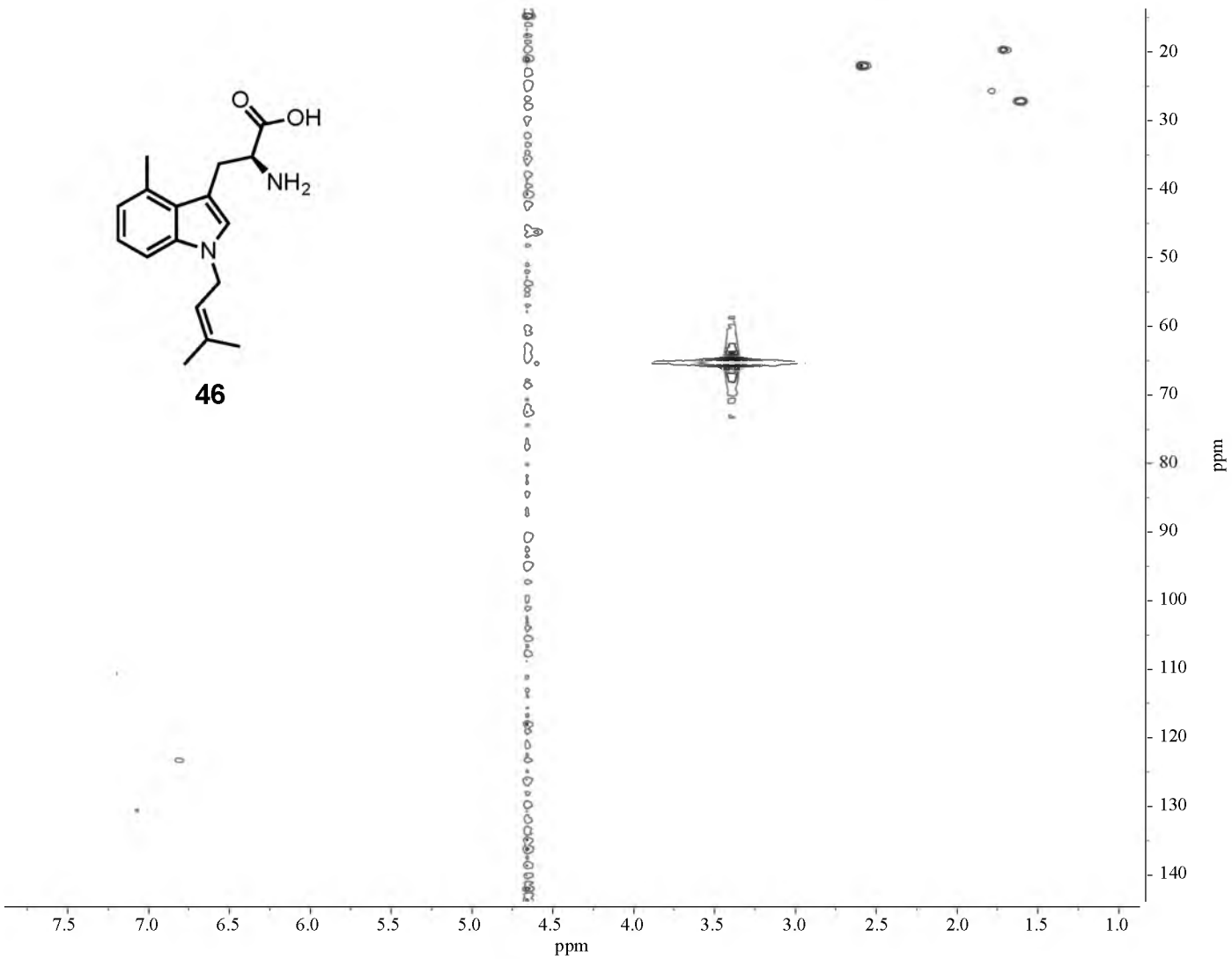
NMR 95. ^{1}H-^{1}H COSY NMR spectrum of compound 46 in D_2O.



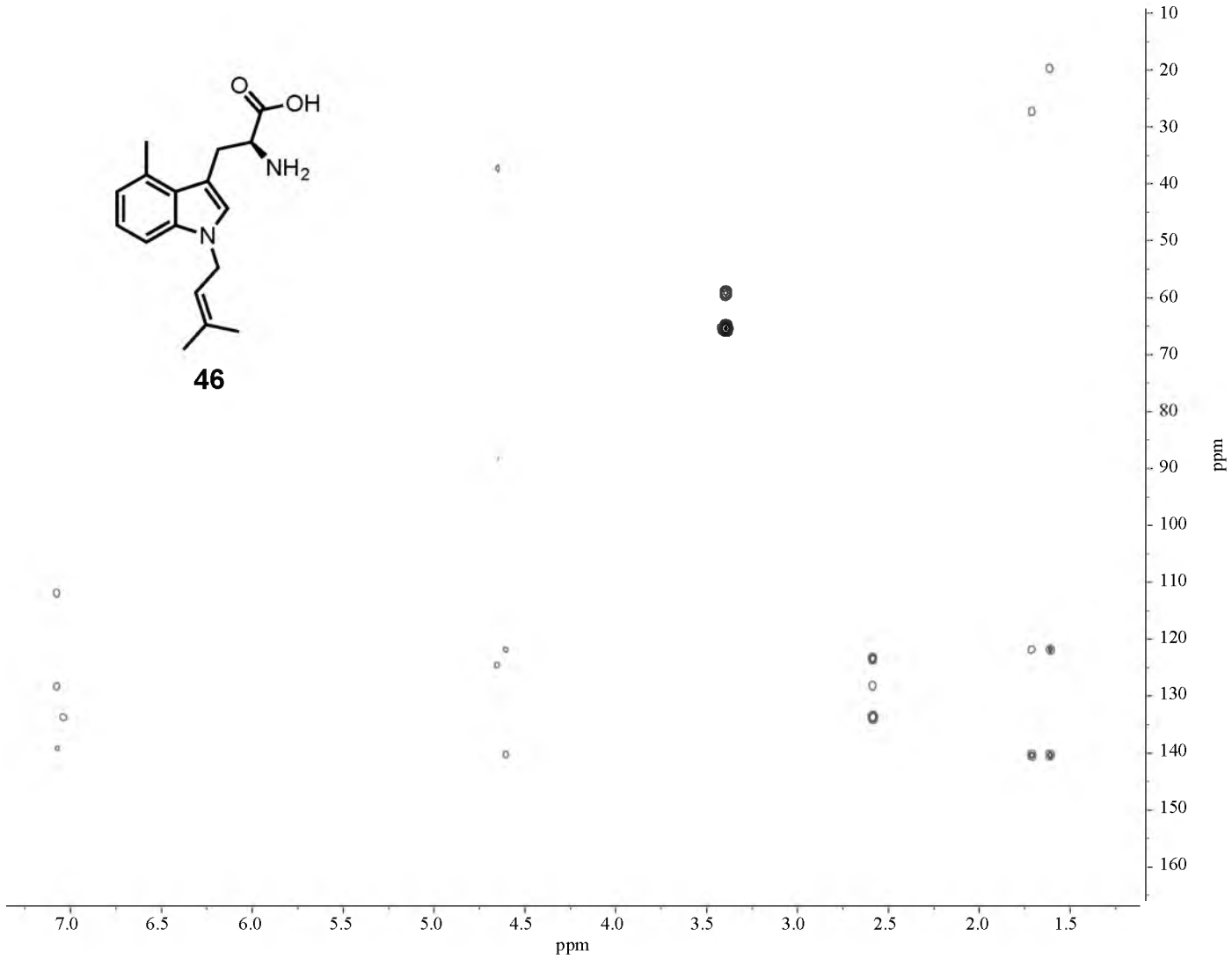
NMR 96. ^{1}H-^{1}H TOCSY NMR spectrum of compound 46 in D_2O.



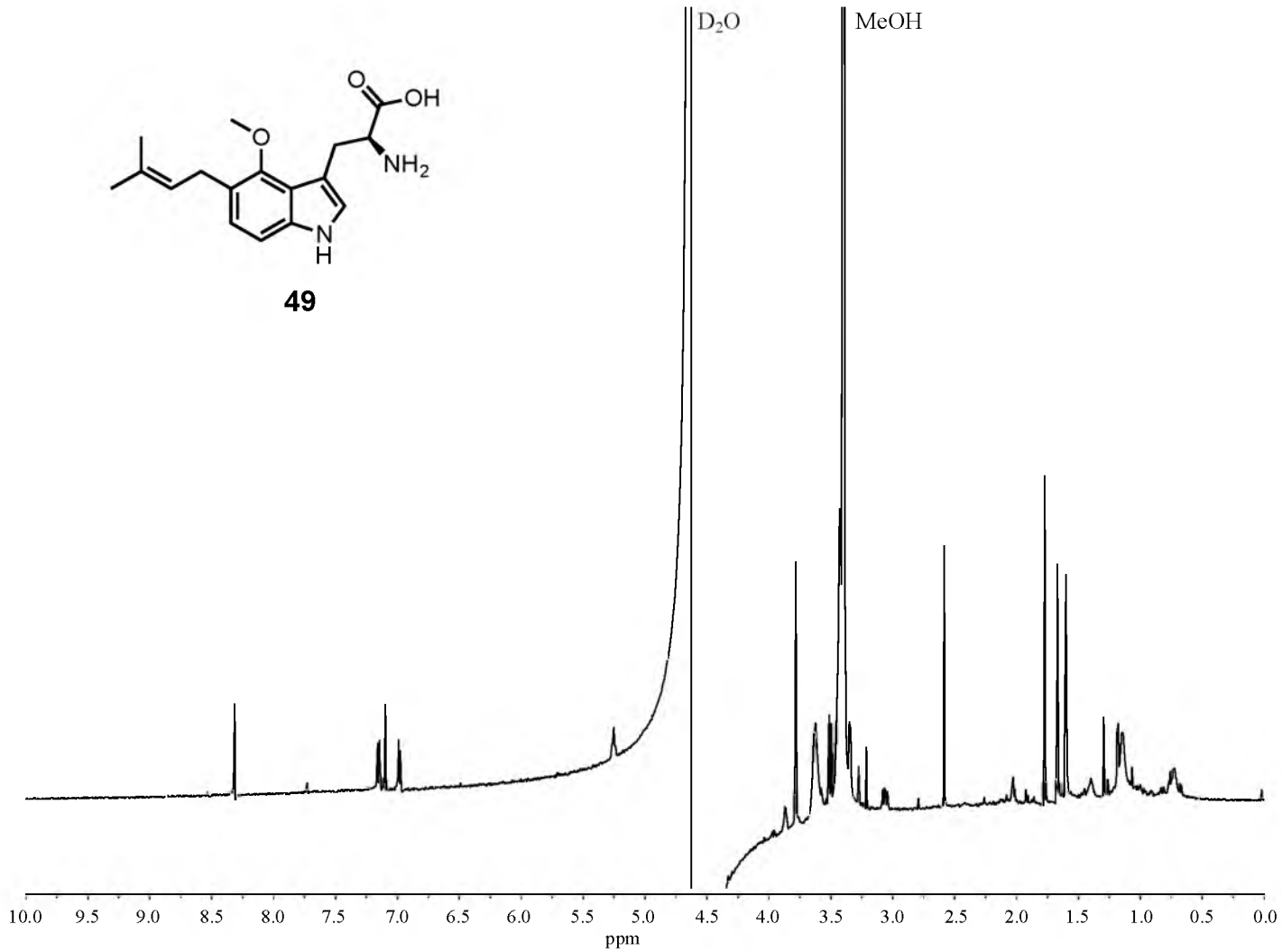
NMR 97. ^{1}H-^{1}H ROESY NMR spectrum of compound 46 in D_2O.



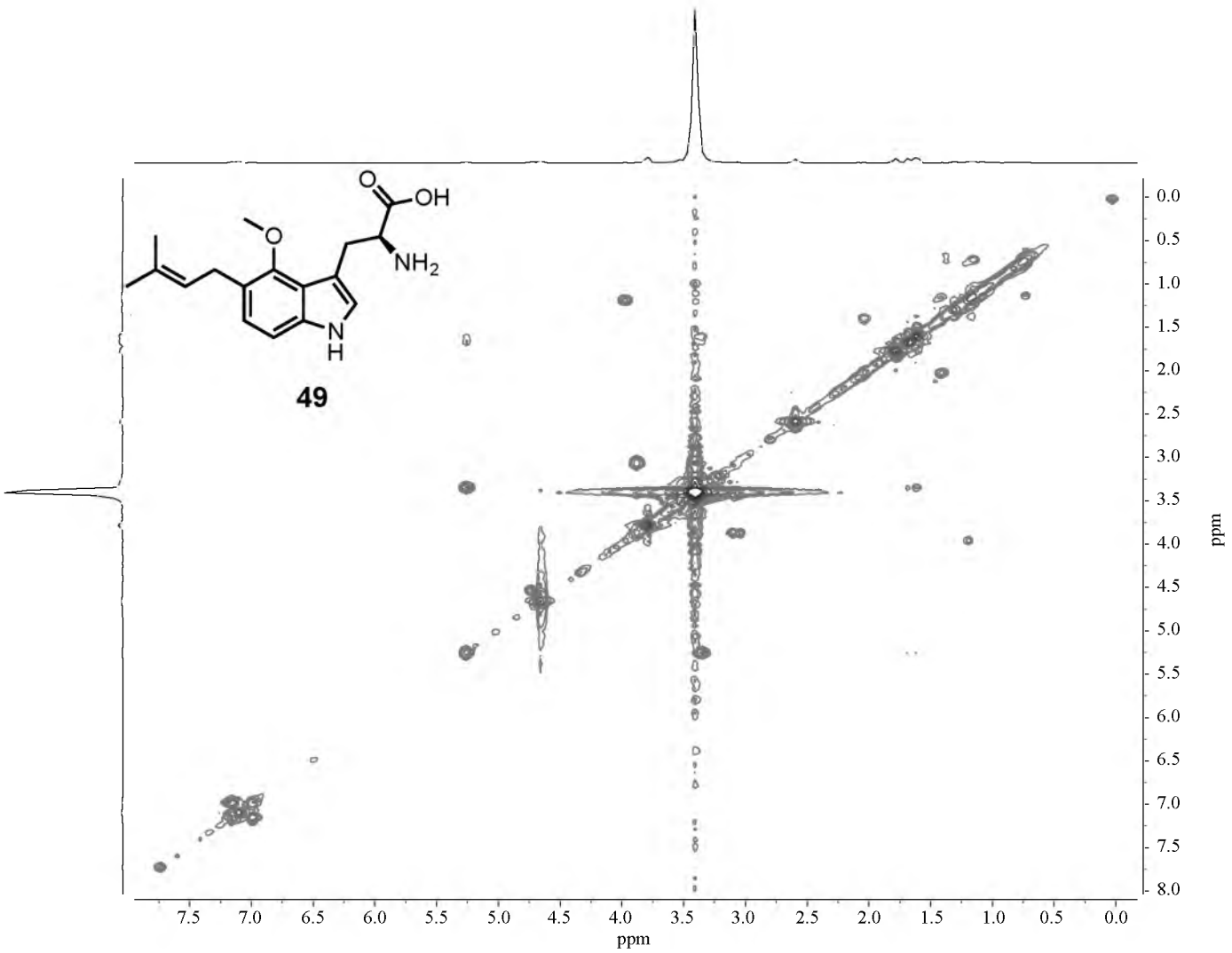
NMR 98. ^{1}H-^{13}C HMQC NMR spectrum of compound 46 in D_2O.



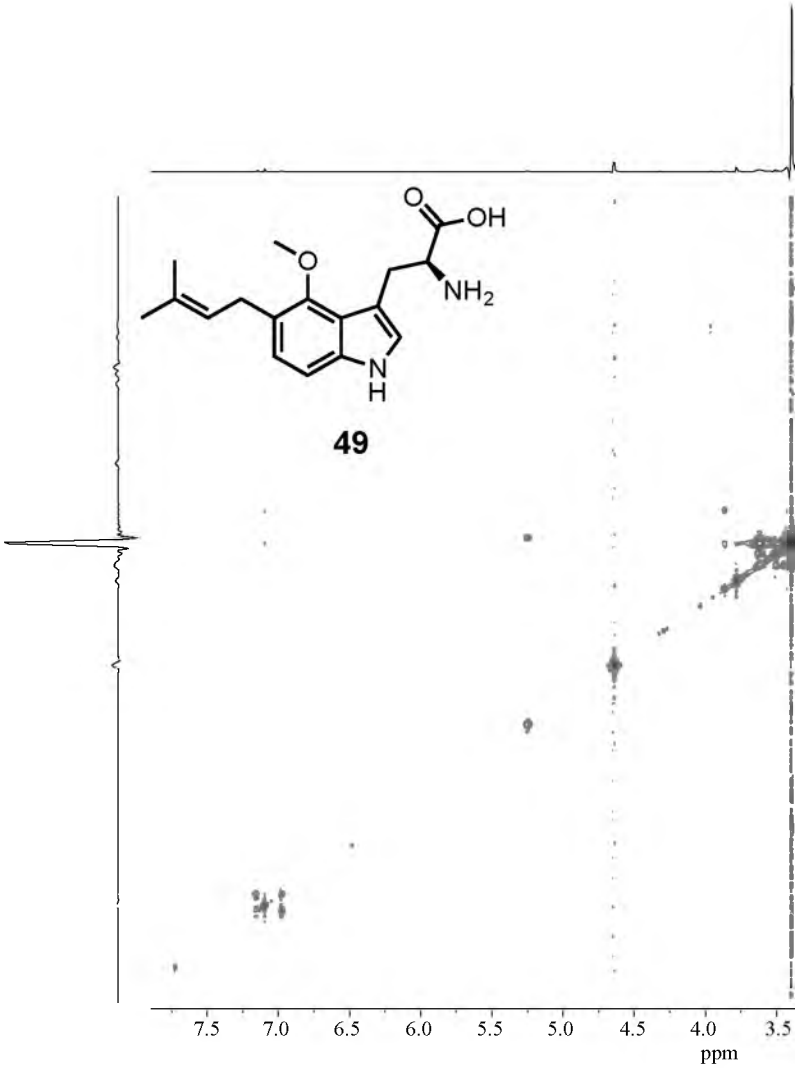
NMR 99. ^{1}H-^{13}C HMBC NMR spectrum of compound 46 in D_2O.



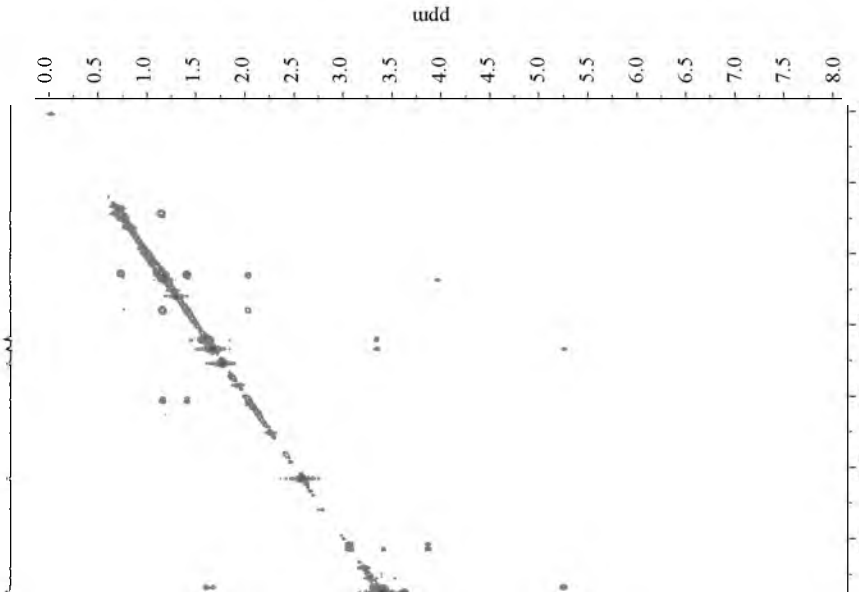
NMR 100. 600 MHz ^{1}H NMR spectrum of compound 49 in D_2O.

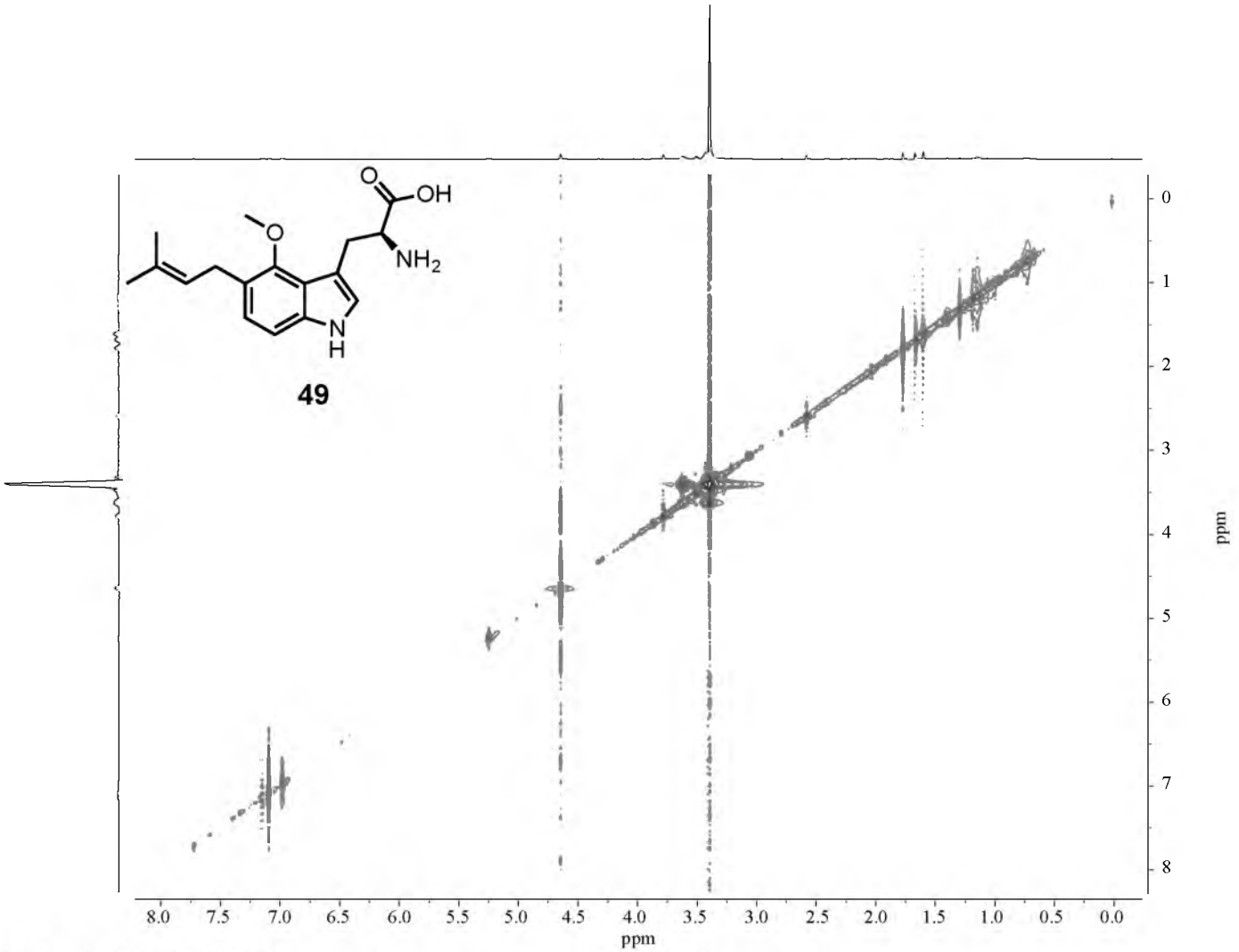


NMR 101. ^{1}H-^{1}H COSY NMR spectrum of compound 49 in D_2O.

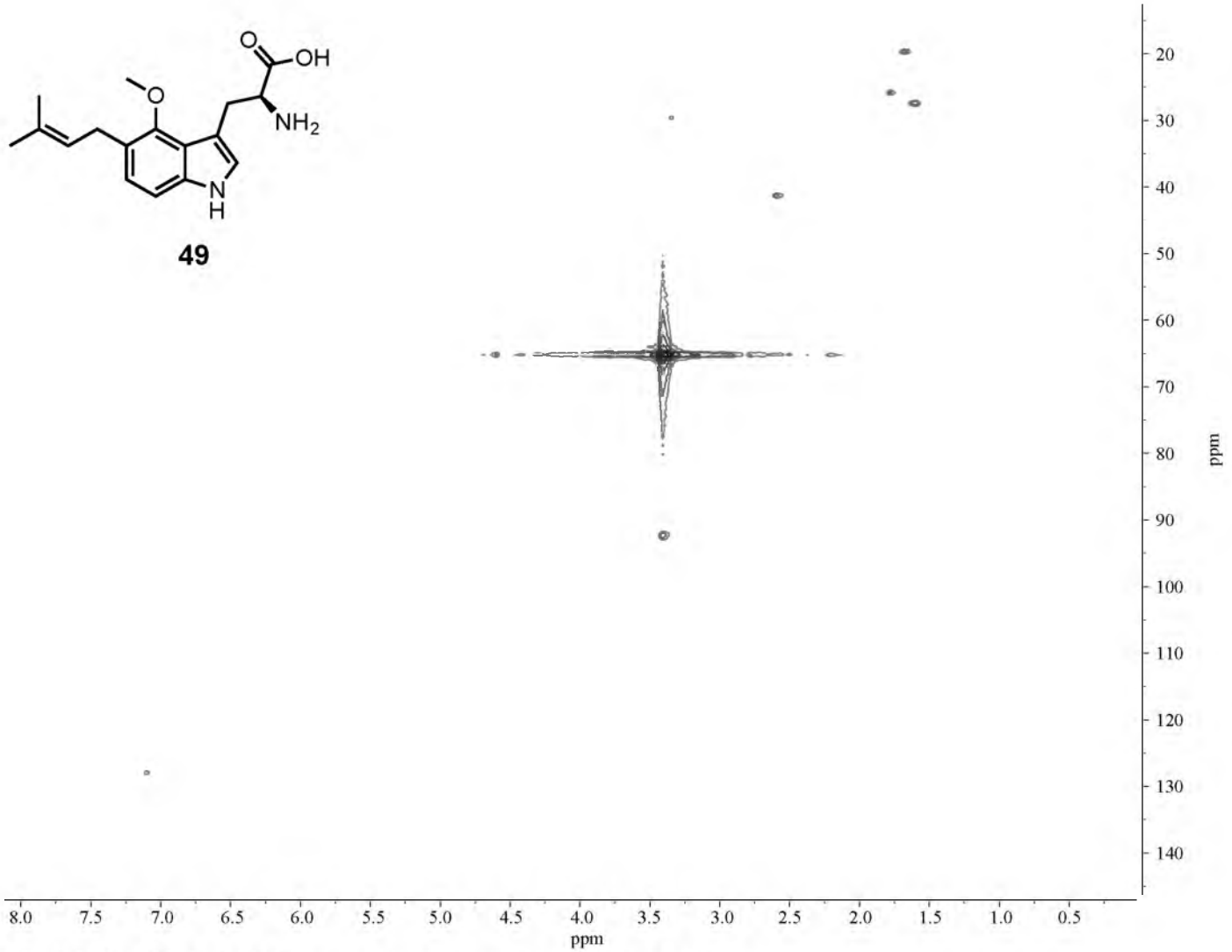


NMR 102. ^{1}H-^{1}H TOCSY NMR spectrum of compound 49 in D_2O.

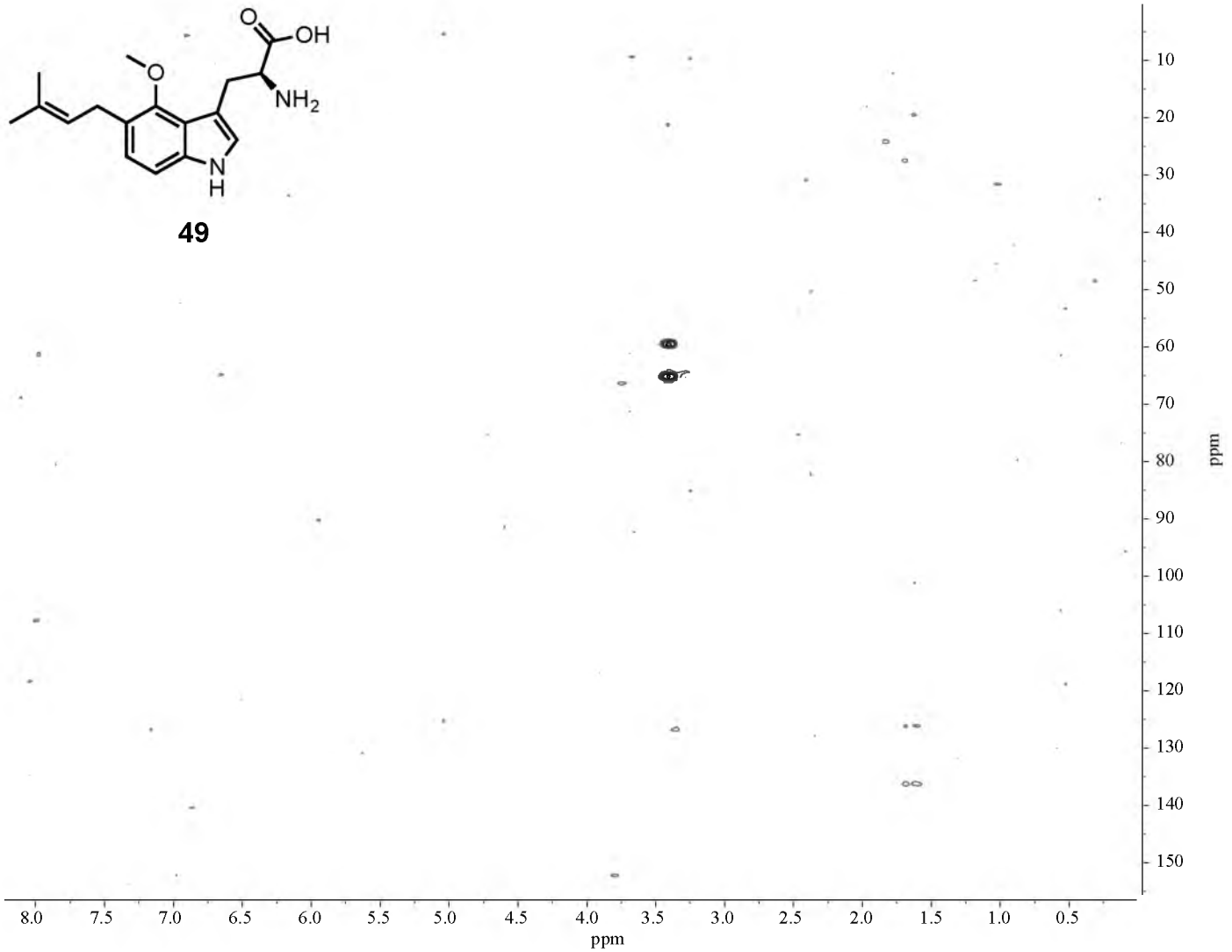

ppm
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0



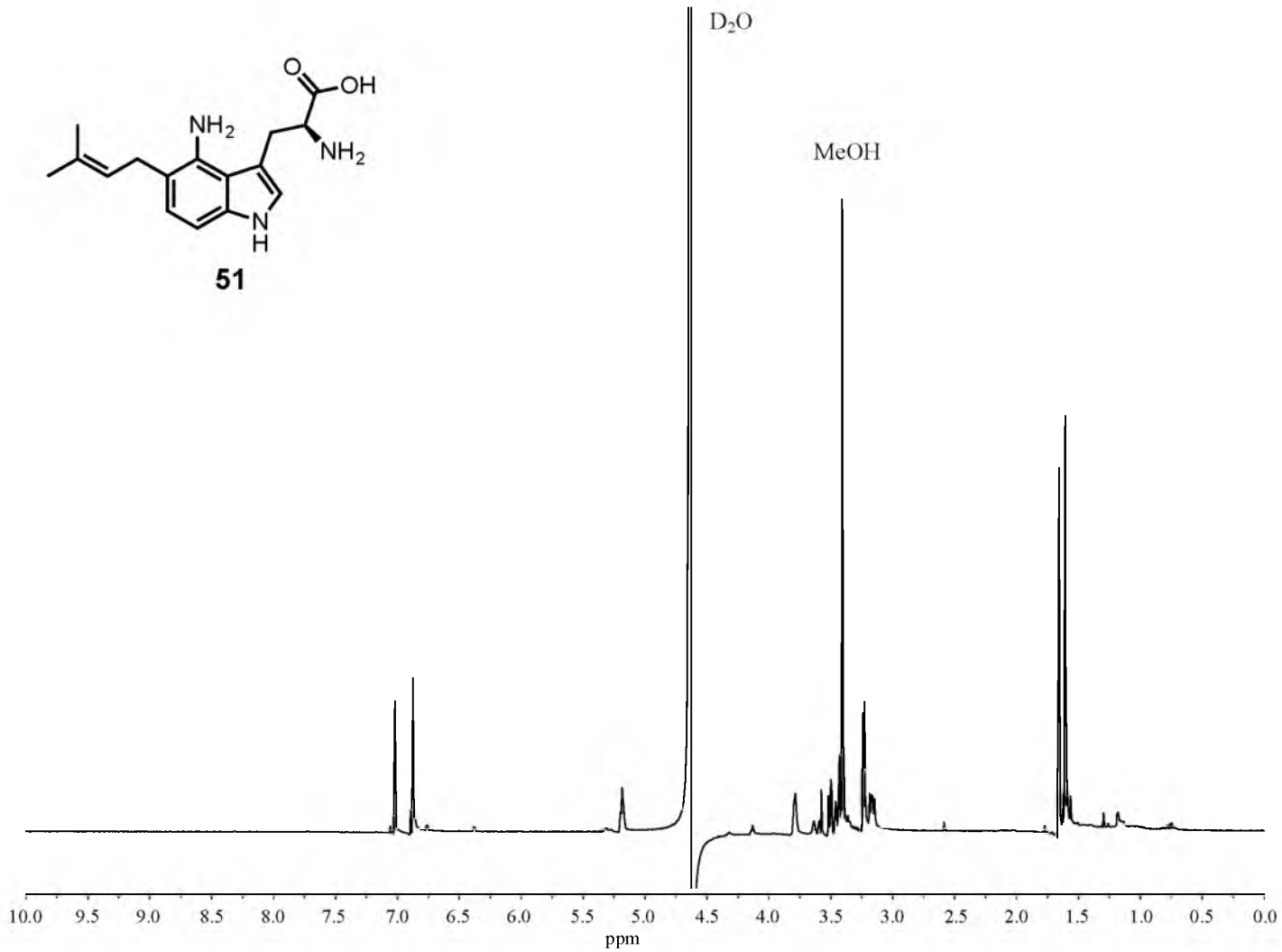
NMR 103. ^{1}H-^{1}H ROESY NMR spectrum of compound 49 in D_2O.



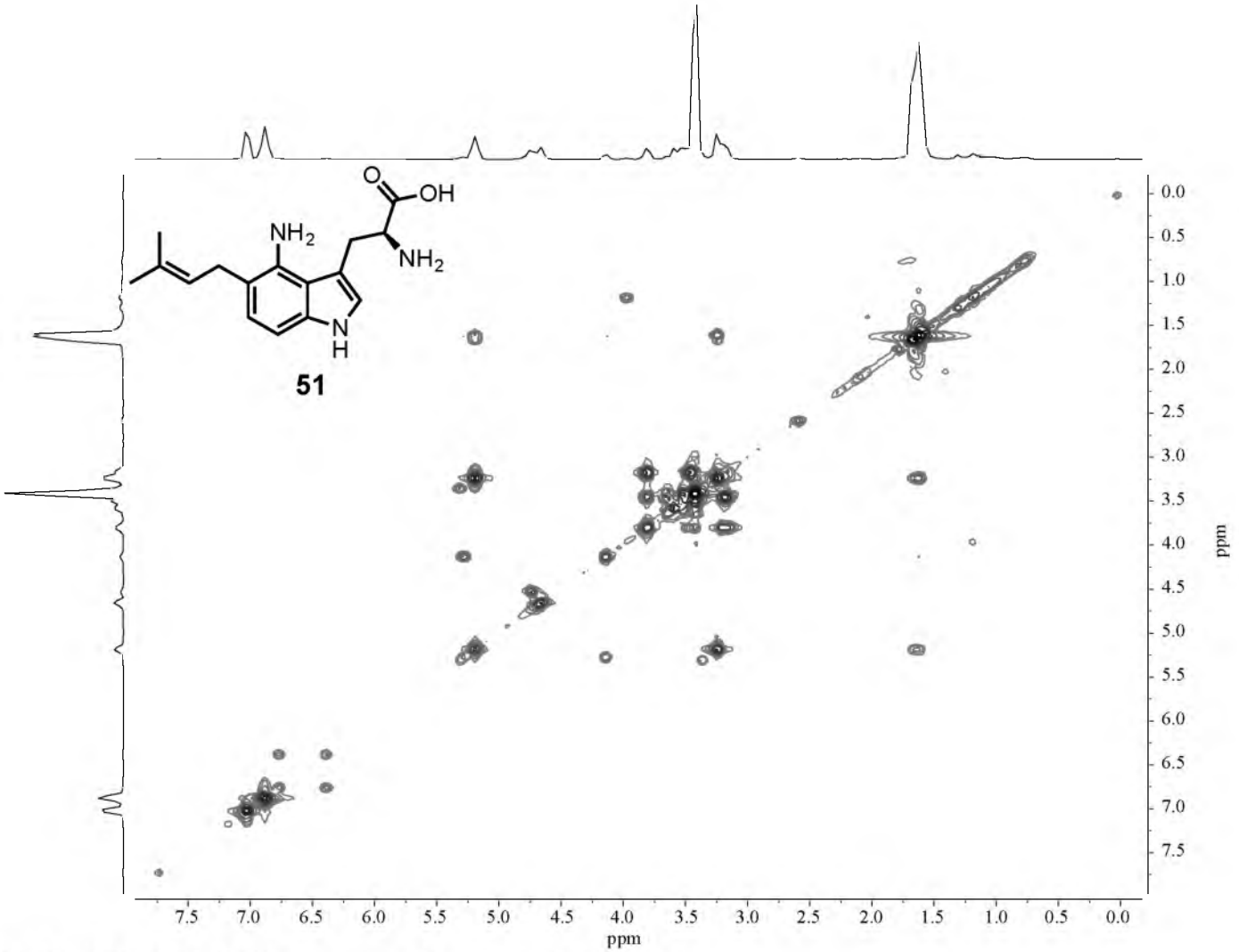
NMR 104. ^{1}H-^{13}C HMQC NMR spectrum of compound 49 in D_2O.



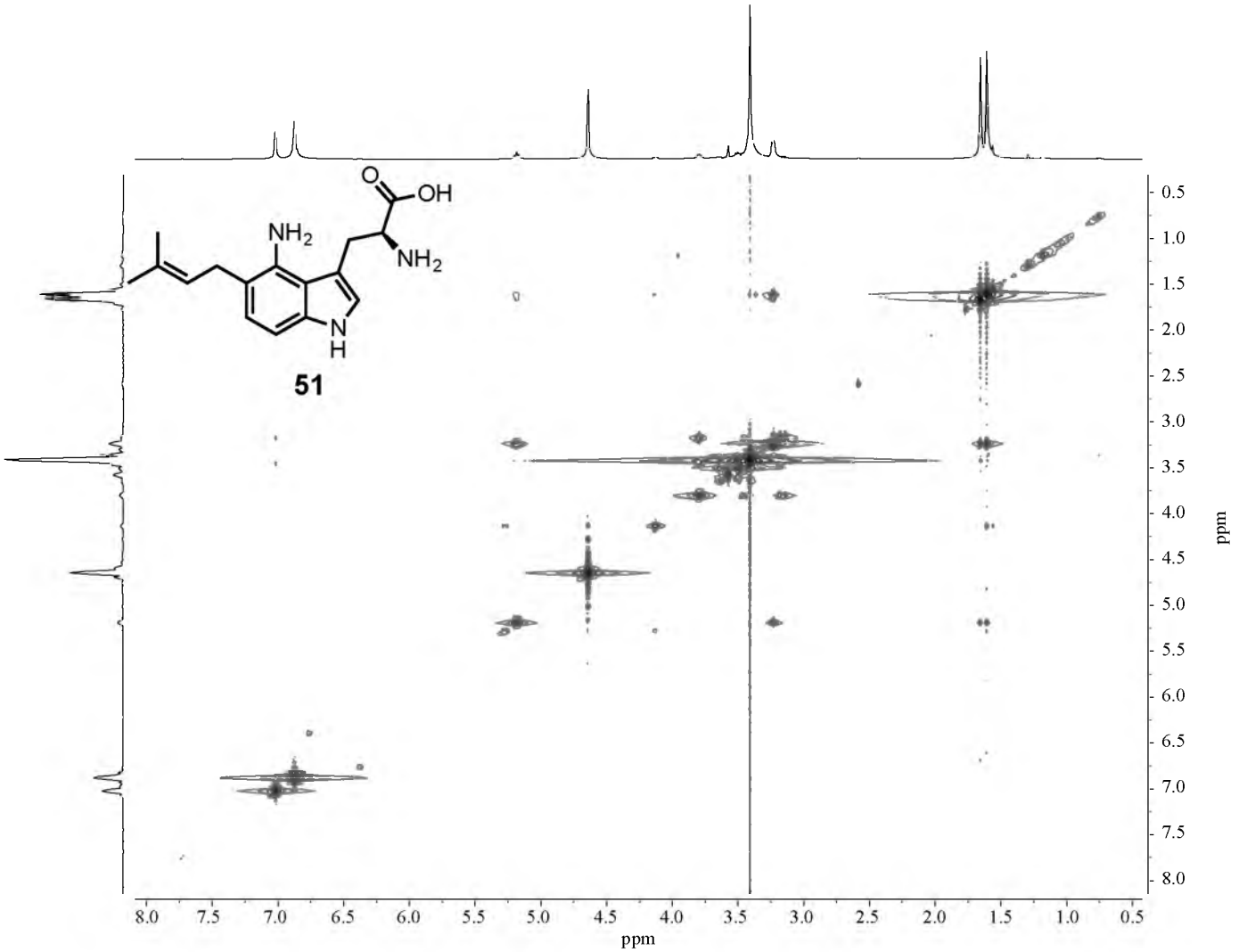
NMR 105. ^{1}H-^{13}C HMBC NMR spectrum of compound 49 in D_2O.



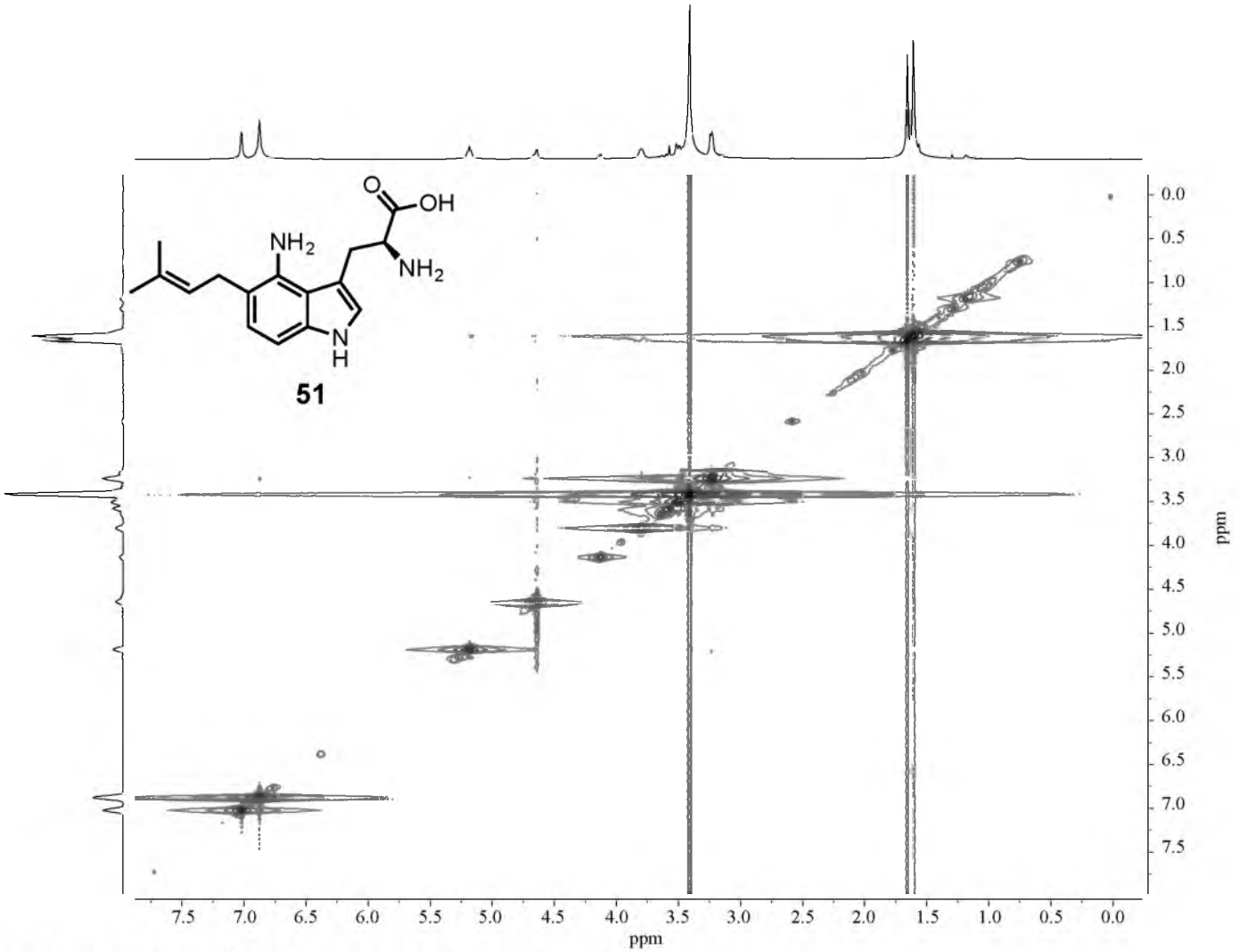
NMR 106. 600 MHz ^{1}H NMR spectrum of compound 51 in D_2O.



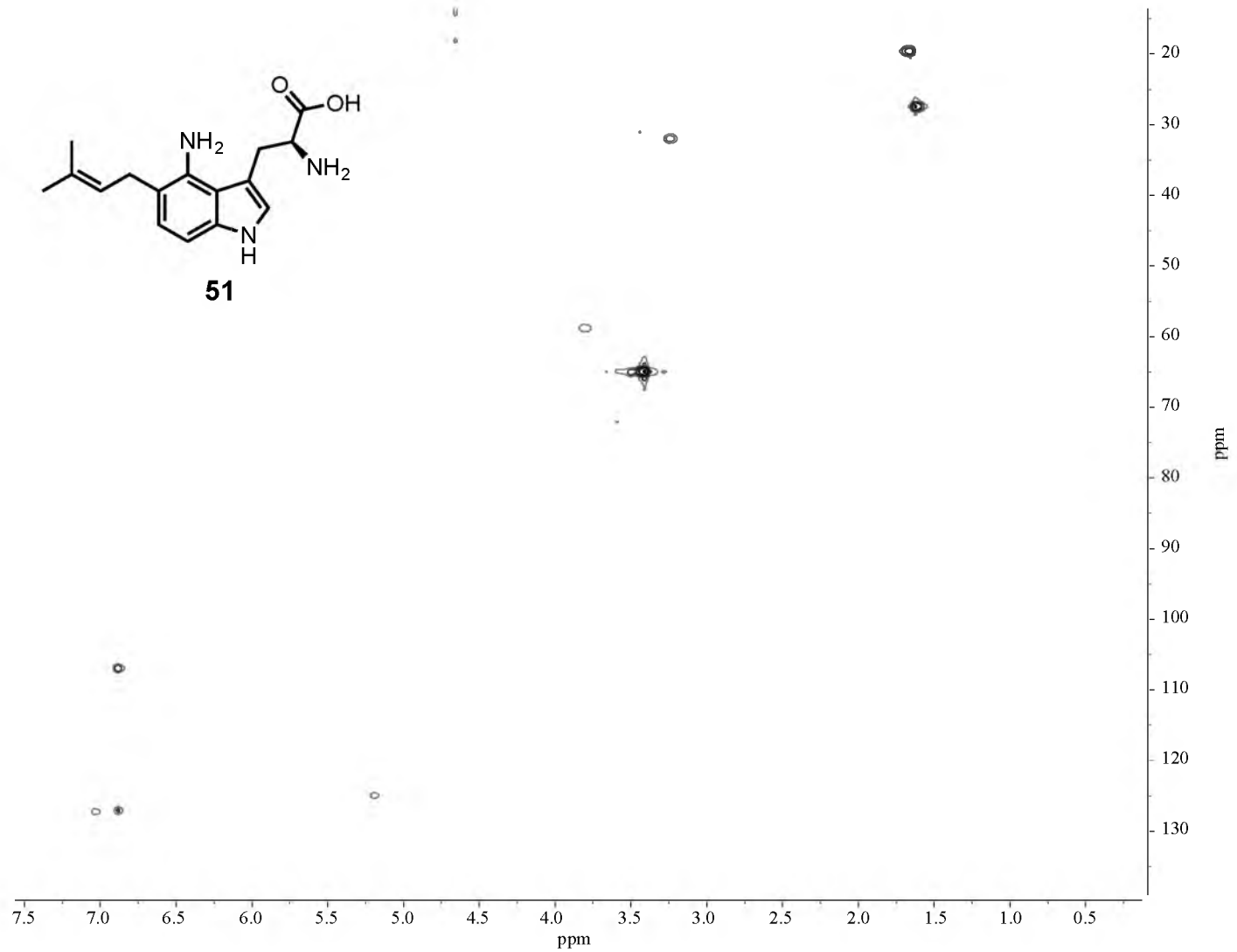
NMR 107. ^{1}H-^{1}H COSY NMR spectrum of compound 51 in D_2O.



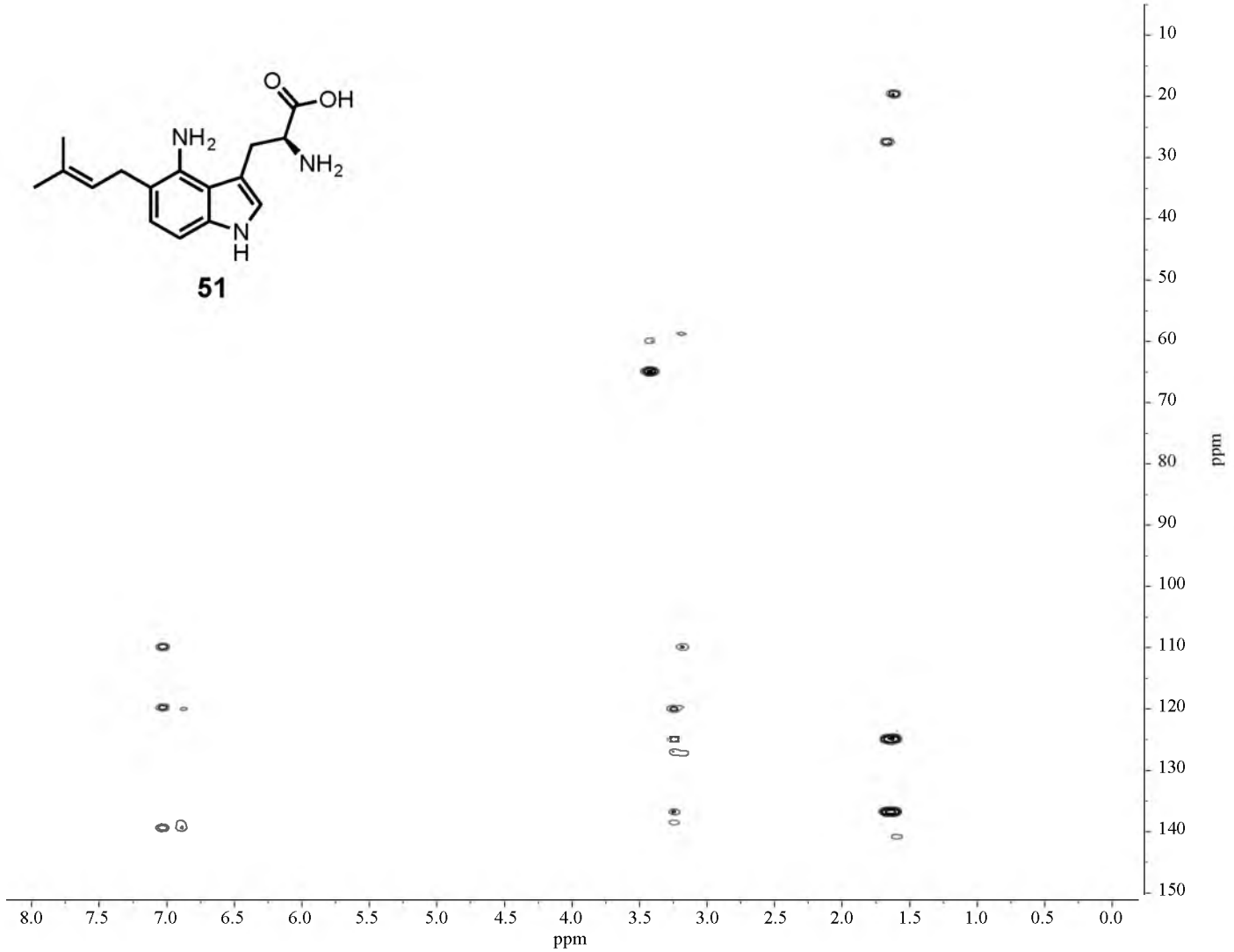
NMR 108. ^{1}H-^{1}H TOCSY NMR spectrum of compound 51 in D_2O.



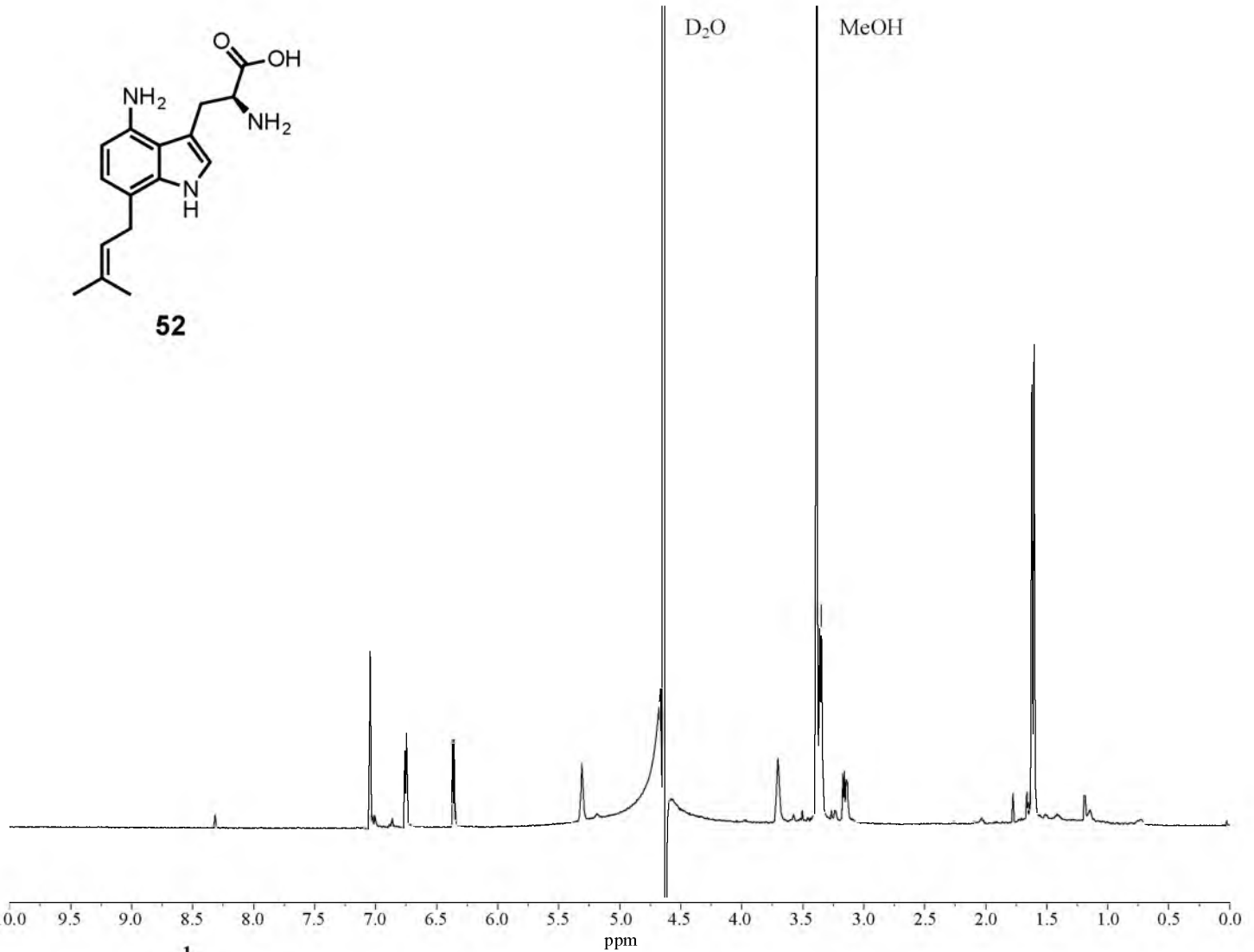
NMR 109. ^{1}H-^{1}H ROESY NMR spectrum of compound 51 in D_2O.



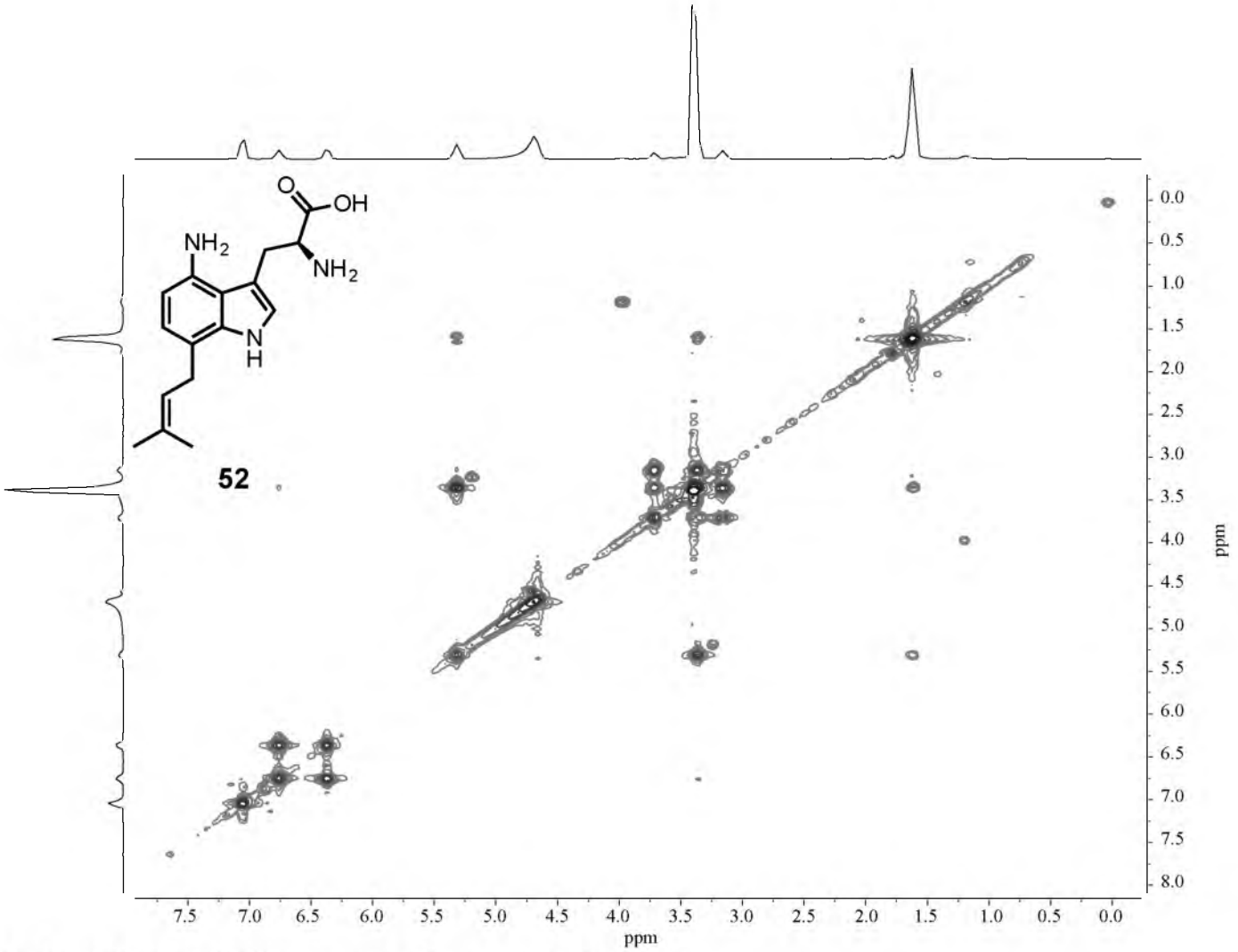
NMR 110. ^{1}H-^{13}C HMQC NMR spectrum of compound 51 in D_2O.



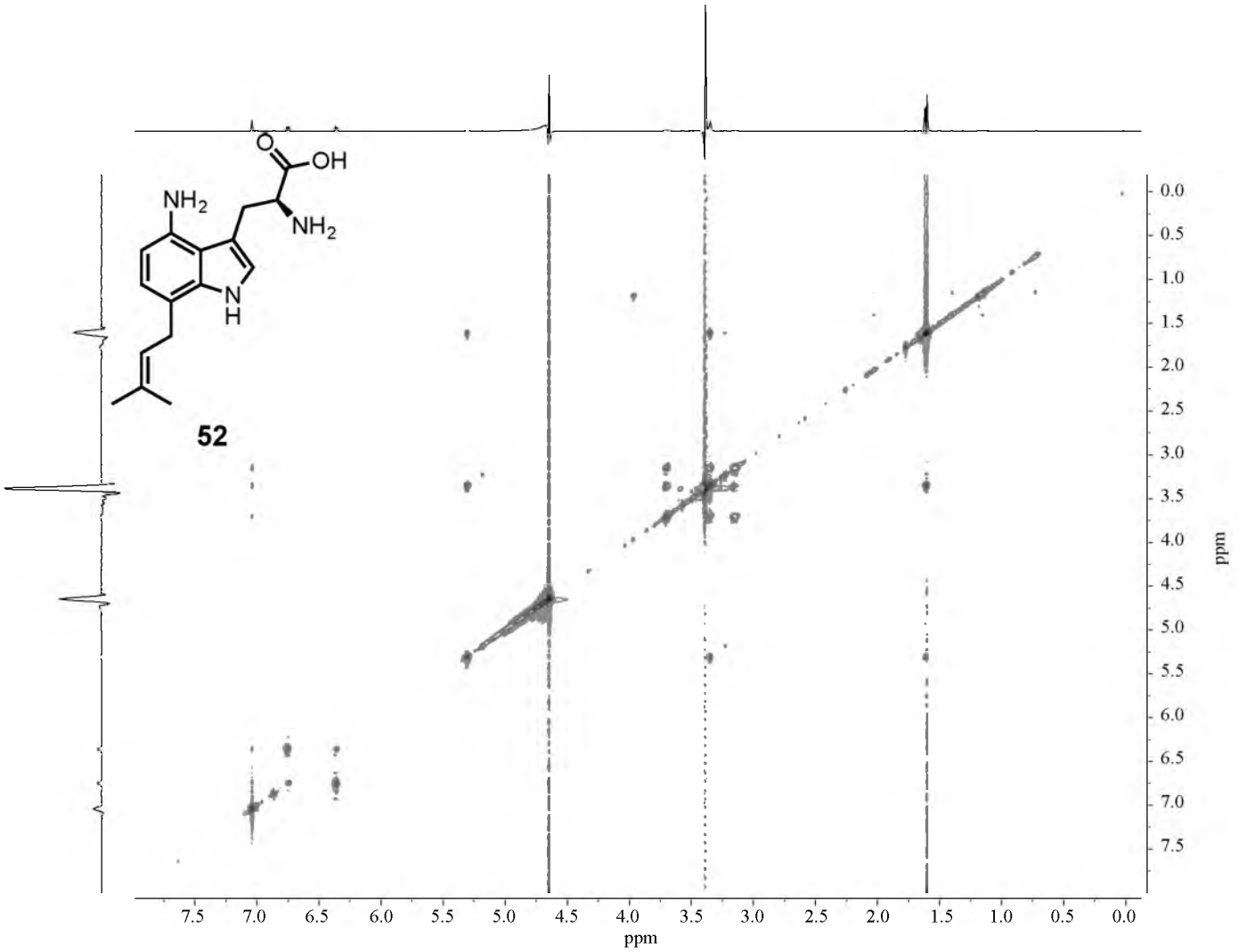
NMR 111. ^{1}H-^{13}C HMBC NMR spectrum of compound 51 in D_2O.



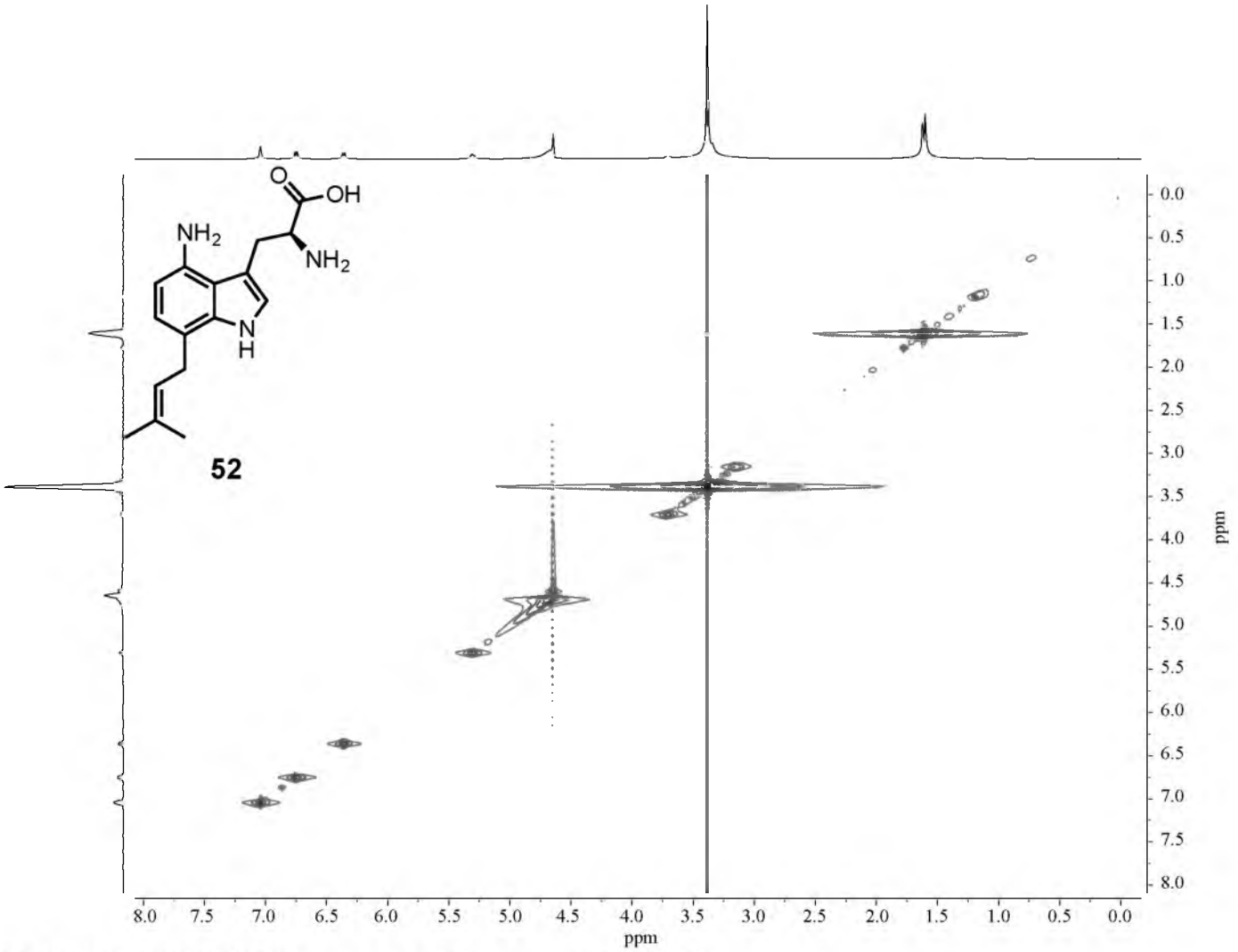
NMR 112. 600 MHz ^{1}H NMR spectrum of compound 52 in D_2O.



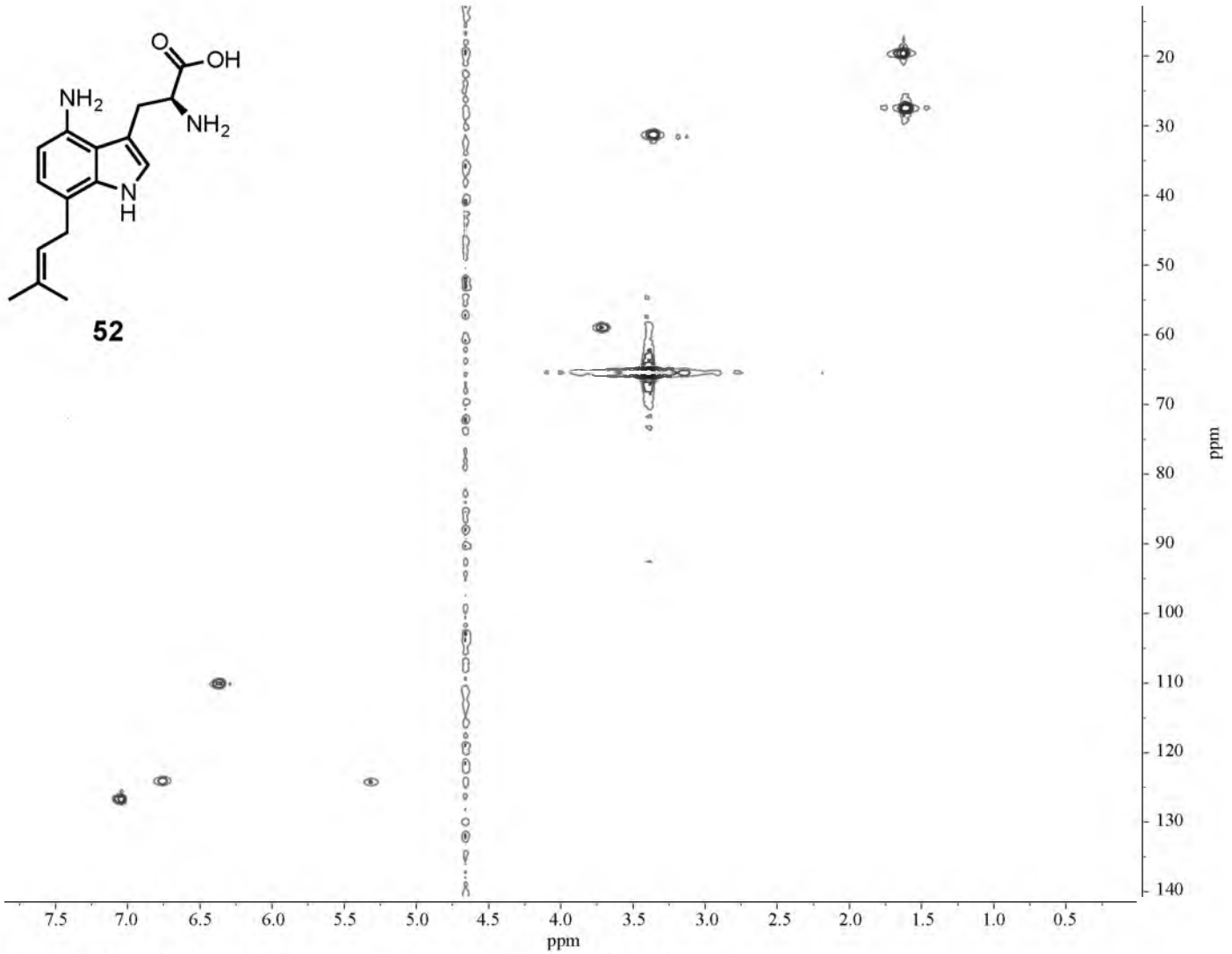
NMR 113. ^{1}H-^{1}H COSY NMR spectrum of compound 52 in D_2O.



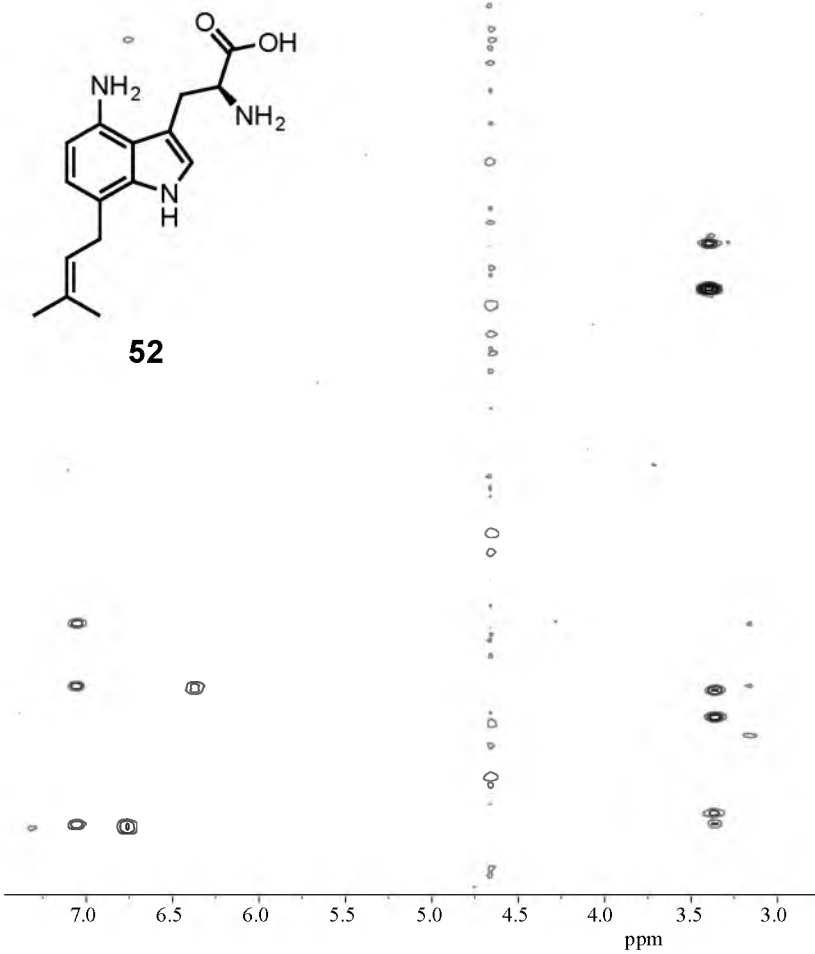
NMR 114. ^{1}H-^{1}H TOCSY NMR spectrum of compound 52 in D_2O.



NMR 115. ^{1}H-^{1}H ROESY NMR spectrum of compound 52 in D_2O.

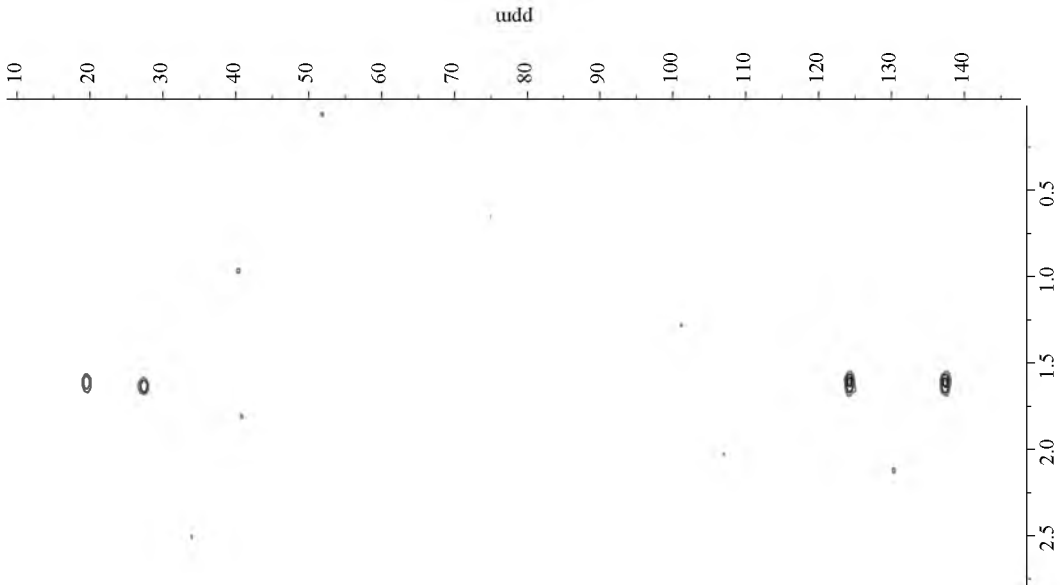


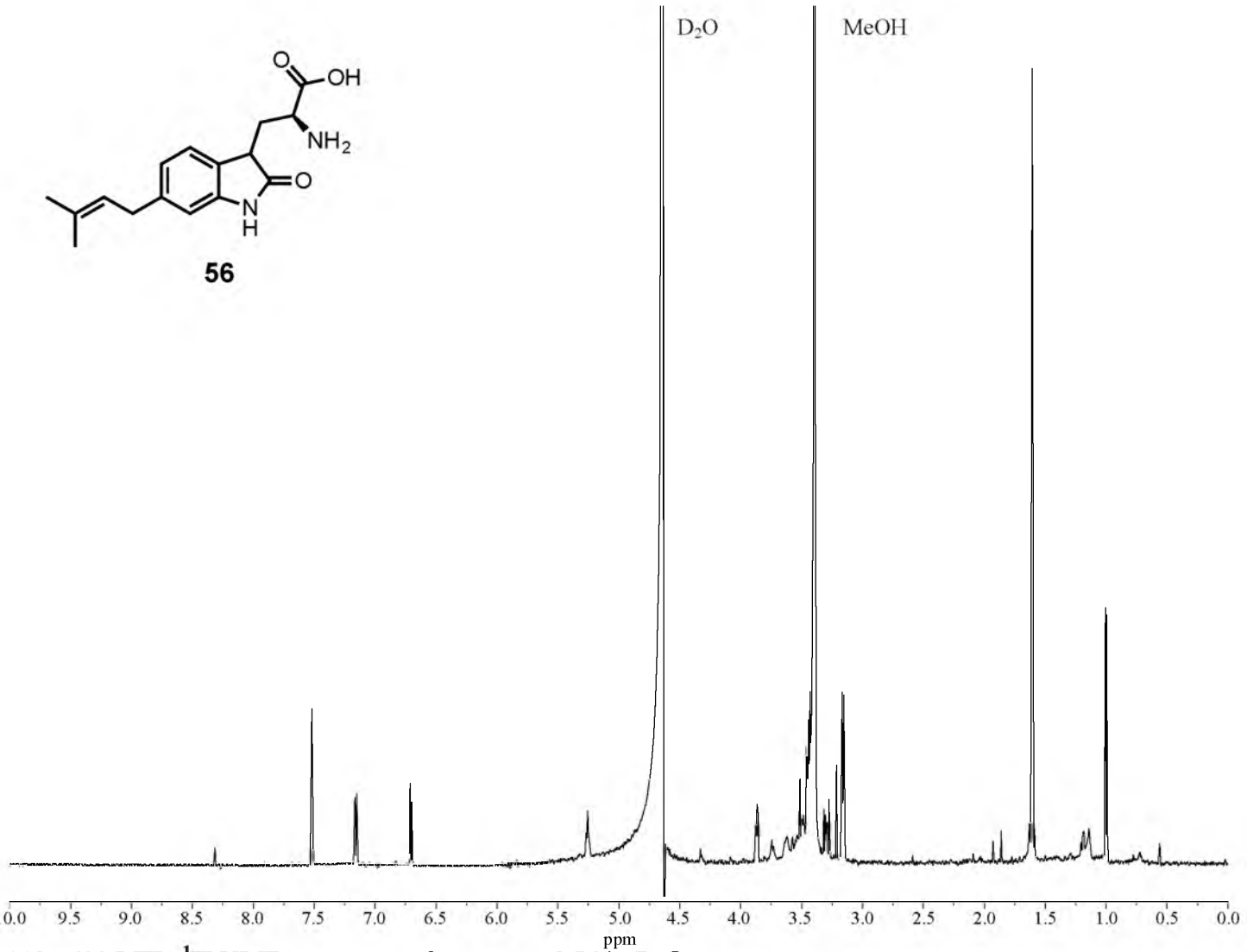
NMR 116. ^{1}H-^{13}C HMQC NMR spectrum of compound 52 in D_2O.



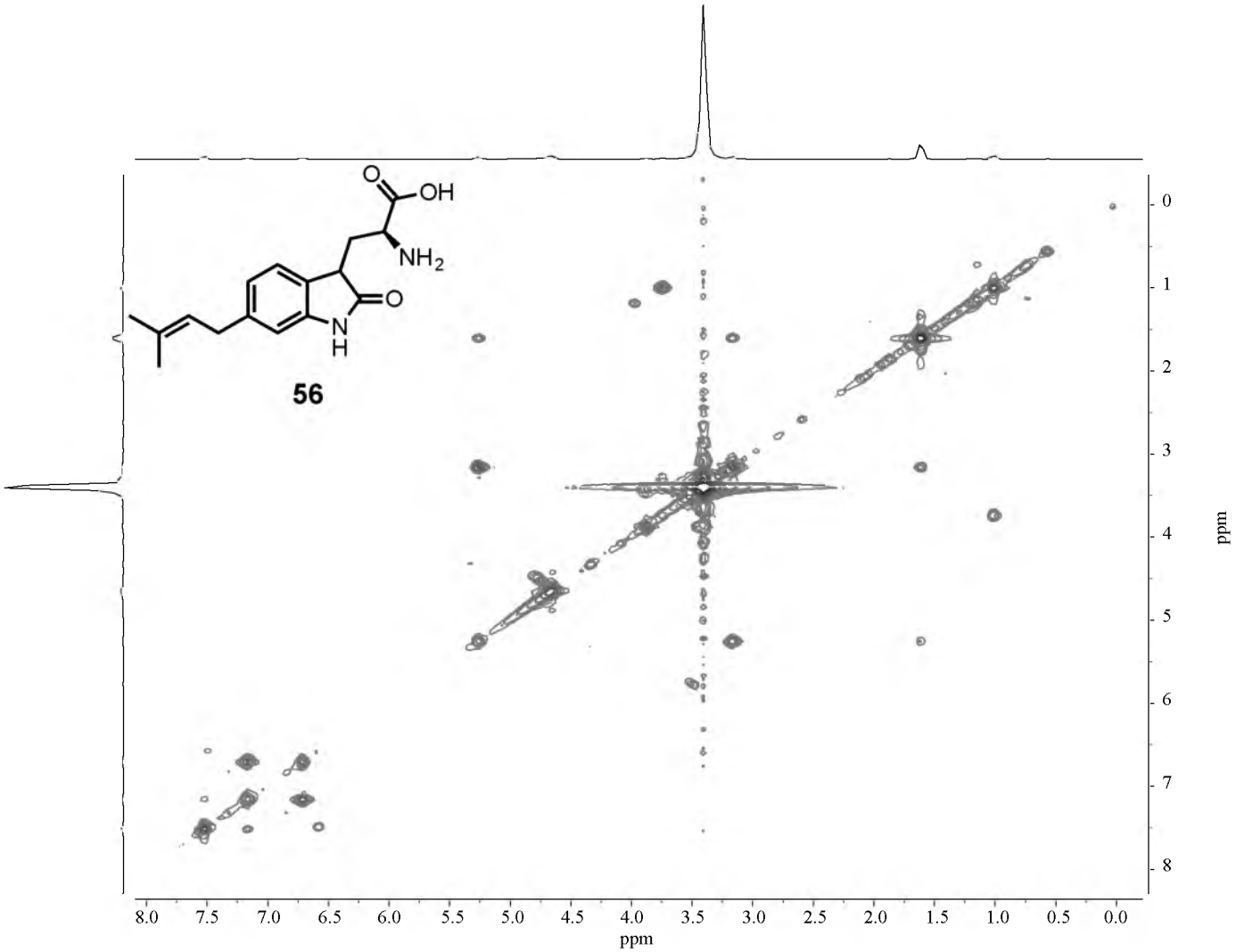
52

NMR 117. ^{1}H-^{13}C HMBC NMR spectrum of compound 52 in D_2O.

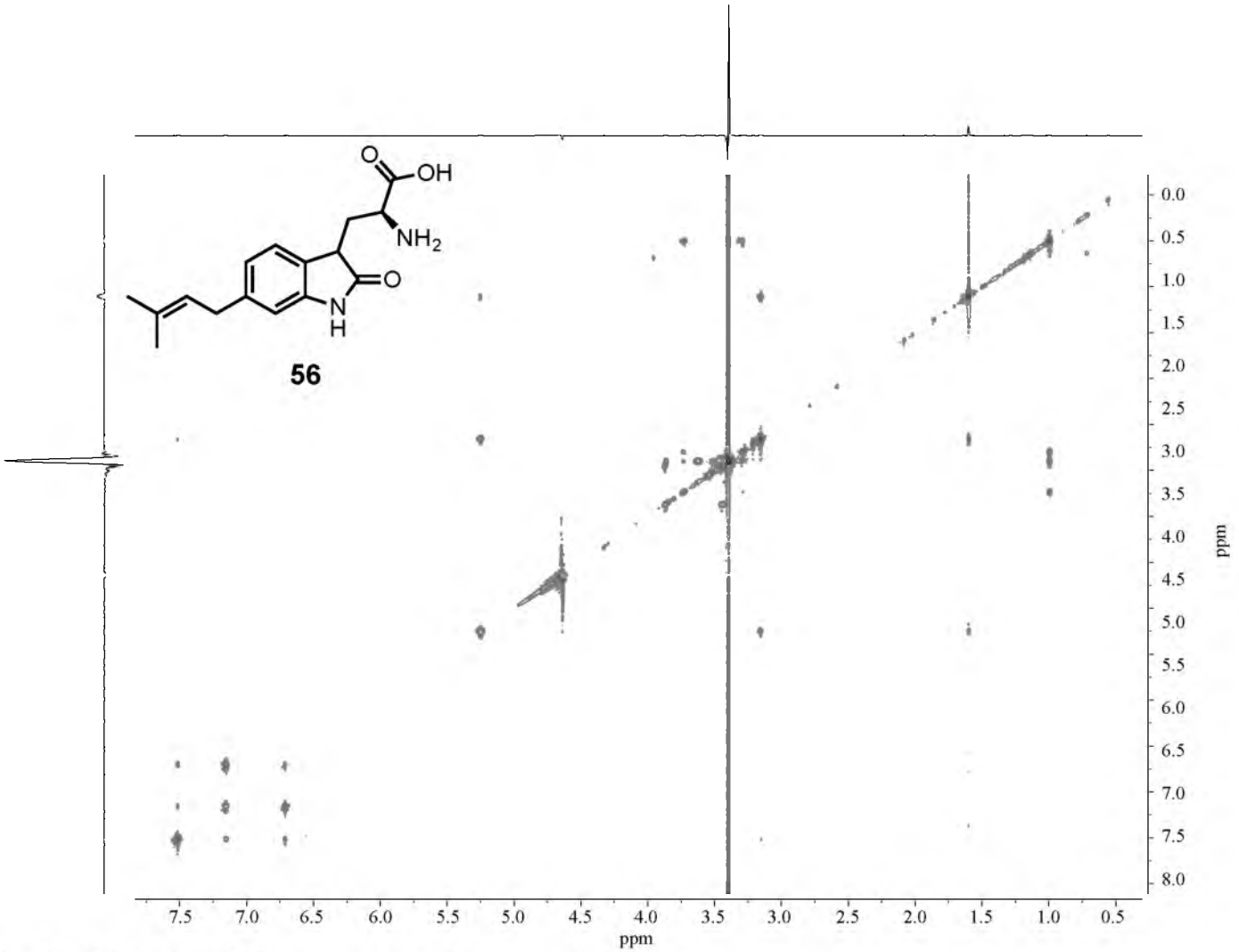

ppm
10
20
30
40
50
60
70
80
90
100
110
120
130
140
0.5
1.0
1.5
2.0
2.5



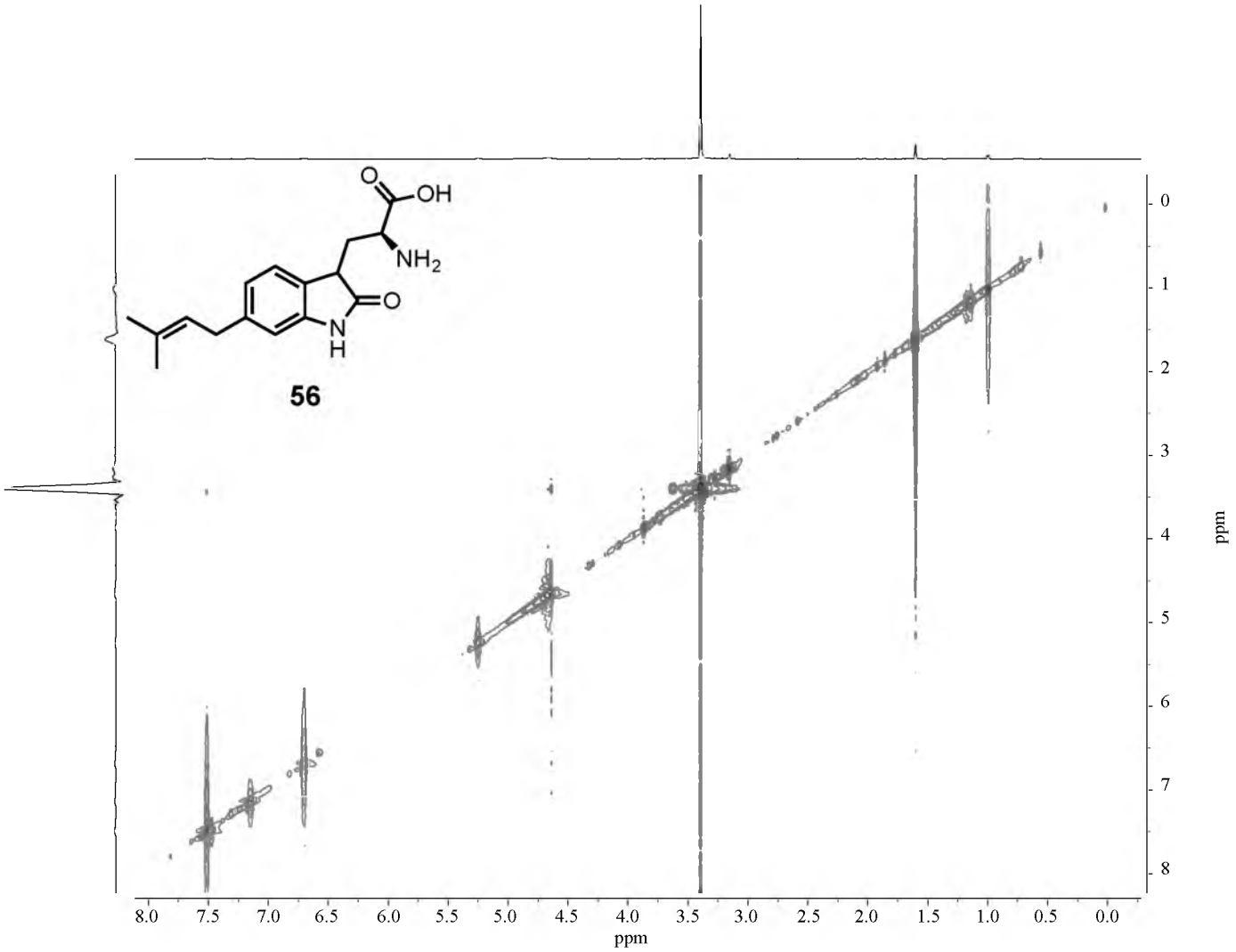
NMR 118. 600 MHz ^{1}H NMR spectrum of compound 56 in D_2O.



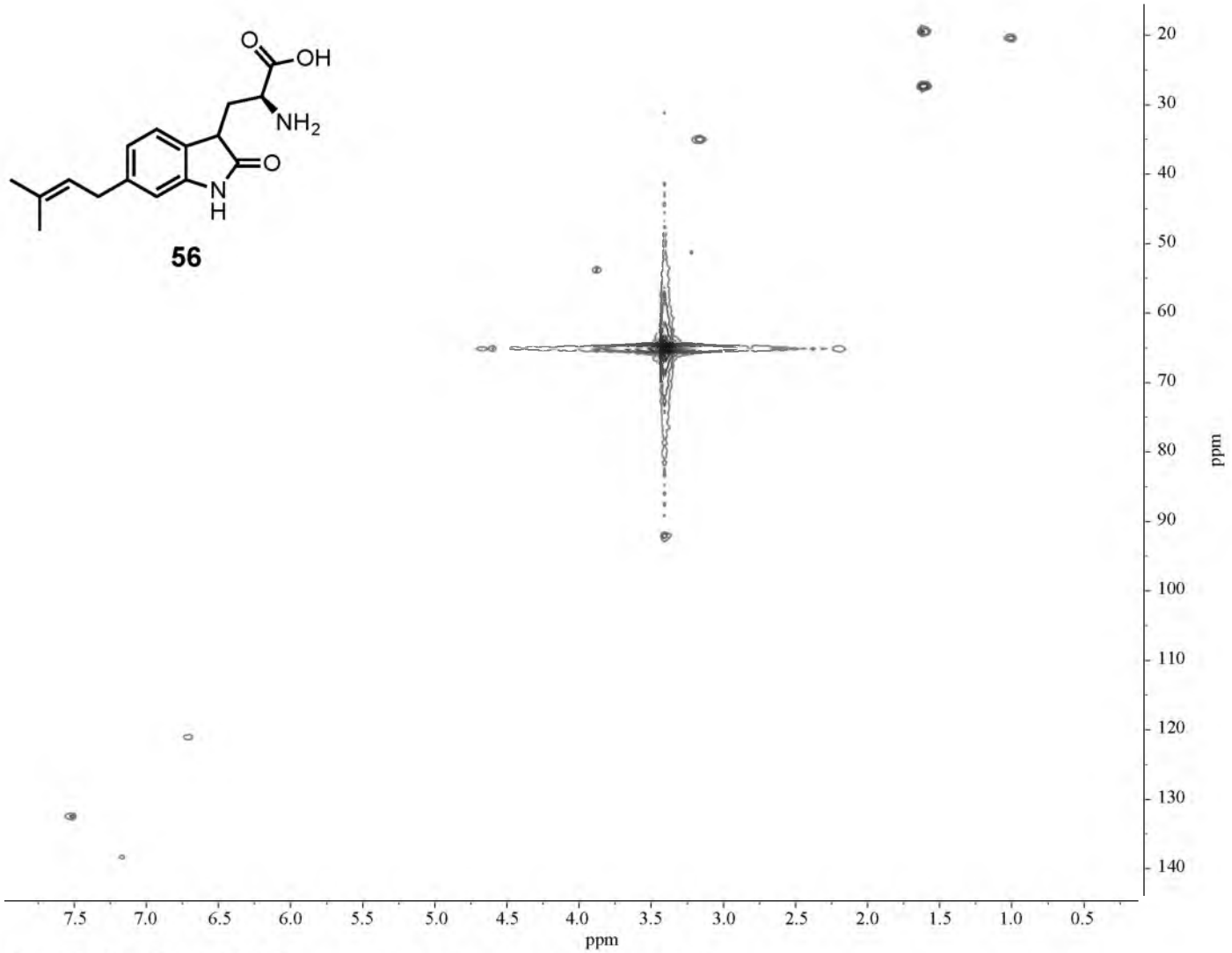
NMR 119. ^{1}H-^{1}H COSY NMR spectrum of compound 56 in D_2O.



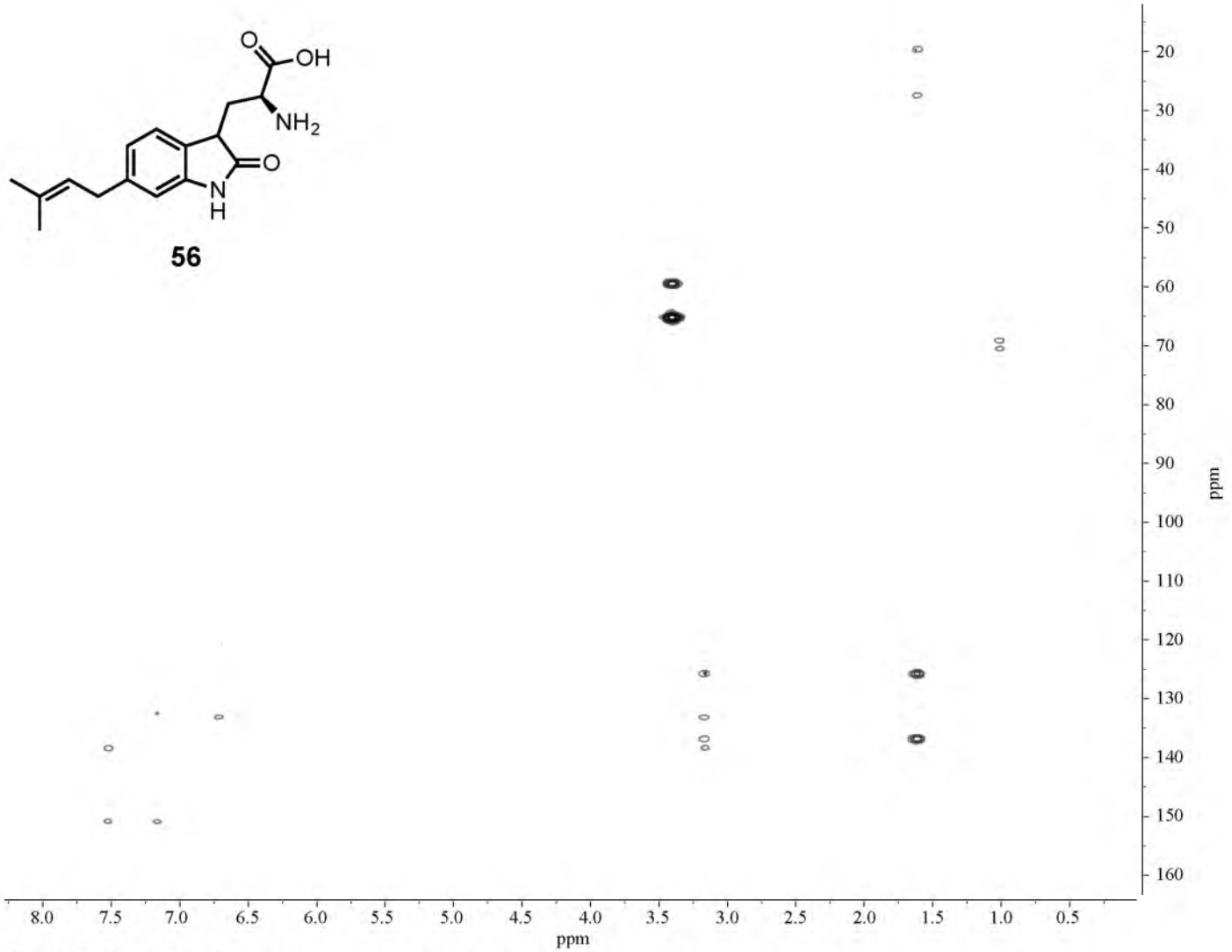
NMR 120. ^{1}H-^{1}H TOCSY NMR spectrum of compound 56 in D_2O.



NMR 121. ^{1}H-^{1}H ROESY NMR spectrum of compound 56 in D_2O.



NMR 122. ^{1}H-^{13}C HMQC NMR spectrum of compound 56 in D_2O.



NMR 123. ^{1}H-^{13}C HMBC NMR spectrum of compound 56 in D_2O.

REFERENCES

(1) The Dictionary of Natural Products Online. http://dnp.chemnetbase.com (accessed November 6, 2012).

(2) Matsumi, R.; Atomi, H.; Driessen, A. J. M.; van der Oost, J. *Res. Microbiol.* **2011**, *162*, 39.

(3) McGarvey, D. J.; Croteau, R. *Plant Cell* **1995**, *7*, 1015.

(4) Chappell, J. *Plant Physiol.* **1995**, *107*, 1.

(5) Heuston, S.; Begley, M.; Gahan, C. G. M.; Hill, C. *Microbiology (Reading, U. K.)* **2012**, *158*, 1389.

(6) Croteau, R.; Kutchan, T. M.; Lewis, N. G. *Biochem. Mol. Biol. Plants* **2000**, 1250.

(7) Schardl, C. L.; Panaccione, D. G.; Tudzynski, P. *Alkaloids (San Diego, CA, U. S.)* **2006**, *63*, 45.

(8) Elliott, C. E.; Gardiner, D. M.; Thomas, G.; Cozijnsen, A.; Van De Wouw, A.; Howlett, B. J. *Mol. Plant Pathol.* **2007**, *8*, 791.

(9) Lichtenthaler, H. K.; Rohmer, M.; Schwender, J. *Physiol. Plant.* **1997**, *101*, 643.

(10) Miziorko, H. M. *Arch. Biochem. Biophys.* **2011**, *505*, 131.

(11) Hunter, W. N. *J. Biol. Chem.* **2007**, *282*, 21573.

(12) Rohmer, M.; Seemann, M.; Horbach, S.; Bringer-Meyer, S.; Sahm, H. *J. Am. Chem. Soc.* **1996**, *118*, 2564.

(13) Poulter, C. D.; Rilling, H. C. *Biosynth. Isoprenoid Compd.* **1981**, *1*, 161.

(14) Poulter, C. D. *Phytochem. Rev.* **2006**, *5*, 17.

(15) Fox, E. M.; Howlett, B. J. *Curr. Opin. Microbiol.* **2008**, *11*, 481.

(16) Hoffmeister, D.; Keller, N. P. *Nat. Prod. Rep.* **2007**, *24*, 393.

(17) Fox, E. M.; Howlett, B. J. *Mycol. Res.* **2008**, *112*, 162.

(18) Gardiner, D. M.; Waring, P.; Howlett, B. J. *Microbiology (Reading, U. K.)* **2005**, *151*, 1021.

(19) Weindling, R.; Emerson, O. H. *Phytopathology* **1936**, *26*, 1068.

(20) Vigushin, D. M.; Mirsaidi, N.; Brooke, G.; Sun, C.; Pace, P.; Inman, L.; Moody, C. J.; Coombes, R. C. *Med. Oncol. (Totowa, NJ, U. S.)* **2004**, *21*, 21.

(21) Howlett, B. J.; Idnurm, A.; Pedras, M. S. C. *Fungal Genet. Biol.* **2001**, *33*, 1.

(22) Li, S.-M. *Nat. Prod. Rep.* **2010**, *27*, 57.

(23) Wallwey, C.; Li, S.-M. *Nat. Prod. Rep.* **2011**, *28*, 496.

(24) Hulvova, H.; Galuszka, P.; Frebortova, J.; Frebort, I. *Biotechnol Adv* **2012**.

(25) Williams, R. M.; Stocking, E. M.; Sanz-Cervera, J. F. *Top. Curr. Chem.* **2000**, *209*, 97.

(26) Seabra, M. C.; Reiss, Y.; Casey, P. J.; Brown, M. S.; Goldstein, J. L. *Cell (Cambridge, Mass.)* **1991**, *65*, 429.

(27) Moore, J. A.; Poulter, C. D. *Biochemistry* **1997**, *36*, 604.

(28) Xie, W.; Zhou, C.; Huang, R. H. *J. Mol. Biol.* **2007**, *367*, 872.

(29) Soderberg, T.; Chen, A.; Poulter, C. D. *Biochemistry* **2001**, *40*, 14847.

(30) Lee, S.-L.; Floss, H. G.; Heinstein, P. *Arch. Biochem. Biophys.* **1976**, *177*, 84.

(31) Liang, P.-H.; Ko, T.-P.; Wang, A. H. J. *Eur. J. Biochem.* **2002**, *269*, 3339.

(32) Zhang, D.; Jennings, S. M.; Robinson, G. W.; Poulter, C. D. *Arch. Biochem. Biophys.* **1991**, *304*, 133.

(33) Clarke, S. *Annu. Rev. Biochem.* **1992**, *61*, 355.

(34) Poulter, C. D.; Rilling, H. C. *Acc. Chem. Res.* **1978**, *11*, 307.

(35) Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871.

(36) Hosfield, D. J.; Zhang, Y.; Dougan, D. R.; Broun, A.; Tari, L. W.; Swanson, R. V.; Finn, J. *J. Biol. Chem.* **2004**, *279*, 8526.

(37) Song, L.; Poulter, C. D. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 3044.

(38) Casey, P. J.; Seabra, M. C. *J. Biol. Chem.* **1996**, *271*, 5289.

(39) Roskoski, R. *Biochem. Biophys. Res. Commun.* **2003**, *303*, 1.

(40) Moores, S. L.; Schaber, M. D.; Mosser, S. D.; Rands, E.; O'Hara, M. B.; Garsky, V. M.; Marshall, M. S.; Pompliano, D. L.; Gibbs, J. B. *J. Biol. Chem.* **1991**, *266*, 14603.

(41) Heide, L. *Current Opinion in Chemical Biology* **2009**, *13*, 171.

(42) Tello, M.; Kuzuyama, T.; Heide, L.; Noel, J. P.; Richard, S. B. *Cell. Mol. Life Sci.* **2008**, *65*, 1459.

(43) Saleh, O.; Haagen, Y.; Seeger, K.; Heide, L. *Phytochemistry (Elsevier)* **2009**, *70*, 1728.

(44) Li, S.-M. *Phytochemistry* **2009**, *70*, 1746.

(45) Suvarna, K.; Stevenson, D.; Meganathan, R.; Hudspeth, M. E. S. *J. Bacteriol.* **1998**, *180*, 2782.

(46) Sadre, R.; Gruber, J.; Frentzen, M. *FEBS Lett.* **2006**, *580*, 5357.

(47) Venkatesh, T. V.; Karunanandaa, B.; Free, D. L.; Rottnek, J. M.; Baszis, S. R.; Valentin, H. E. *Planta* **2006**, *223*, 1134.

(48) Melzer, M.; Heide, L. *Biochim. Biophys. Acta, Lipids Lipid Metab.* **1994**, *1212*, 93.

(49) Forsgren, M.; Attersand, A.; Lake, S.; Gruenler, J.; Swiezewska, E.; Dallner, G.; Climent, I. *Biochem. J.* **2004**, *382*, 519.

(50) Braeuer, L.; Brandt, W.; Schulze, D.; Zakharova, S.; Wessjohann, L. *ChemBioChem* **2008**, *9*, 982.

(51) Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. *Science* **1997**, *277*, 1815.

(52) Yazaki, K.; Sasaki, K.; Tsurumaru, Y. *Phytochemistry* **2009**, *70*, 1739.

(53) Botta, B.; Delle Monache, G.; Menendez, P.; Boffi, A. *Trends Pharmacol. Sci.* **2005**, *26*, 606.

(54) Yazaki, K.; Kunihisa, M.; Fujisaki, T.; Sato, F. *J. Biol. Chem.* **2002**, *277*, 6240.

(55) Pojer, F.; Wemakor, E.; Kammerer, B.; Chen, H.; Walsh, C. T.; Li, S.-M.; Heide, L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 2316.

(56) Kuzuyama, T.; Noel, J. P.; Richard, S. B. *Nature (London, U. K.)* **2005**, *435*, 983.

(57) Metzger, U.; Keller, S.; Stevenson, C. E. M.; Heide, L.; Lawson, D. M. *Journal of Molecular Biology* **2010**, *404*, 611.

(58) Haagen, Y.; Unsoeld, I.; Westrich, L.; Gust, B.; Richard, S. B.; Noel, J. P.; Heide, L. *FEBS Lett.* **2007**, *581*, 2889.

(59) Metzger, U.; Schall, C.; Zocher, G.; Unsoeld, I.; Stec, E.; Li, S.-M.; Heide, L.; Stehle, T. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 14309.

(60) Schuller, J. M.; Zocher, G.; Liebhold, M.; Xie, X.; Stahl, M.; Li, S.-M.; Stehle, T. *J. Mol. Biol.* **2012**, *422*, 87.

(61) Jost, M.; Zocher, G.; Tarcz, S.; Matuschek, M.; Xie, X.; Li, S.-M.; Stehle, T. *Journal of the American Chemical Society* **2010**, *132*, 17849.

(62) Tsai, H.-F.; Wang, H.; Gebler, J. C.; Poulter, C. D.; Schardl, C. L. *Biochem. Biophys. Res. Commun.* **1995**, *216*, 119.

(63) Gebler, J. C.; Poulter, C. D. *Arch. Biochem. Biophys.* **1992**, *296*, 308.

(64) Gardiner, D. M.; Cozijnsen, A. J.; Wilson, L. M.; Pedras, M. S. C.; Howlett, B. J. *Mol. Microbiol.* **2004**, *53*, 1307.

(65) Ferezou, J. P.; Quesneau-Thierry, A.; Servy, C.; Zissmann, E.; Barbier, M. *J. Chem. Soc., Perkin Trans. 1* **1980**, 1739.

(66) Bu'Lock, J. D.; Clough, L. E. *Aust. J. Chem.* **1992**, *45*, 39.

(67) Pedras, M. S. C. *Recent Res. Dev. Phytochem.* **2001**, *5*, 109.

(68) Pedras, M. S. C.; Yu, Y. *Can. J. Chem.* **2009**, *87*, 556.

(69) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.

(70) Grundmann, A.; Kuznetsova, T.; Afiyatullov, S. S.; Li, S.-M. *ChemBioChem* **2008**, *9*, 2059.

(71) Grundmann, A.; Li, S.-M. *Microbiology (Reading, U. K.)* **2005**, *151*, 2199.

(72) Yin, W.-B.; Grundmann, A.; Cheng, J.; Li, S.-M. *J. Biol. Chem.* **2009**, *284*, 100.

(73) Yin, W.-B.; Yu, X.; Xie, X.-L.; Li, S.-M. *Org. Biomol. Chem.* **2010**, *8*, 2430.

(74) Yu, X.; Liu, Y.; Xie, X.; Zheng, X.-D.; Li, S.-M. *J. Biol. Chem.* **2012**, *287*, 1371.

(75) Takahashi, S.; Takagi, H.; Toyoda, A.; Uramoto, M.; Nogawa, T.; Ueki, M.; Sakaki, Y.; Osada, H. *J. Bacteriol.* **2010**, *192*, 2839.

(76) Kremer, A.; Westrich, L.; Li, S.-M. *Microbiology (Reading, U. K.)* **2007**, *153*, 3409.

(77) Zou, H.-X.; Xie, X.-L.; Linne, U.; Zheng, X.-D.; Li, S.-M. *Org. Biomol. Chem.* **2010**, *8*, 3037.

(78) Unsoeld, I. A.; Li, S.-M. *Microbiology (Reading, U. K.)* **2005**, *151*, 1499.

(79) Ma, J.; Zuo, D.; Song, Y.; Wang, B.; Huang, H.; Yao, Y.; Li, W.; Zhang, S.; Zhang, C.; Ju, J. *ChemBioChem* **2012**, *13*, 547.

(80) Gebler, J. C.; Woodside, A. B.; Poulter, C. D. *J. Am. Chem. Soc.* **1992**, *114*, 7354.

(81) Steffan, N.; Unsoeld, I. A.; Li, S.-M. *ChemBioChem* **2007**, *8*, 1298.

(82) Luk, L. Y. P.; Tanner, M. E. *J. Am. Chem. Soc.* **2009**, *131*, 13932.

(83) Cress, W. A.; Chayet, L. T.; Rilling, H. C. *J. Biol. Chem.* **1981**, *256*, 10917.

(84) Wang, H., University of Utah, 1998.

(85) Sundberg, R. J. *The Chemistry of Indoles (Organic Chemistry; a Series of Monographs, Vol. 18)*; Academic Press: New York, 1970.

(86) Shibuya, M.; Chou, H. M.; Fountoulakis, M.; Hassam, S.; Kim, S. U.; Kobayashi, K.; Otsuka, H.; Rogalska, E.; Cassady, J. M.; Floss, H. G. *J. Am. Chem. Soc.* **1990**, *112*, 297.

(87) Zocher, G.; Saleh, O.; Heim Joel, B.; Herbst Dominik, A.; Heide, L.; Stehle, T. *PLoS One* **2012**, *7*, e48427.

(88) Phan, R. M.; Poulter, C. D. *J. Org. Chem.* **2001**, *66*, 6705.

(89) Luk, L. Y. P.; Qian, Q.; Tanner, M. E. *J. Am. Chem. Soc.* **2011**, *133*, 12342.

(90) Escher, E.; Bernier, M.; Parent, P. *Helv. Chim. Acta* **1983**, *66*, 1355.

(91) Lu, H. S. M.; Volk, M.; Kholodenko, Y.; Gooding, E.; Hochstrasser, R. M.; DeGrado, W. F. *Journal of the American Chemical Society* **1997**, *119*, 7173.

(92) Yao, Z.-J.; Gao, Y.; Burke, T. R., Jr. *Tetrahedron: Asymmetry* **1999**, *10*, 3727.

(93) Fischer, E. *Ber.* **1886**, *19*, 1563.

(94) Reissert, A. *Ber. Dtsch. chem. Ges.* **1897**, *30*, 1030.

(95) Batcho, A. D.; Leimgruber, W. *Org. Synth.* **1985**, *63*, 214.

(96) Sugasawa, T.; Adachi, M.; Sasakura, K.; Kitagawa, A. *J. Org. Chem.* **1979**, *44*, 578.

(97) Albertson, N. F.; Archer, S.; Suter, C. M. *J. Am. Chem. Soc.* **1945**, *67*, 36.

(98) Gebler, J., University of Utah, 1990.

(99) Chenault, H. K.; Dahmer, J.; Whitesides, G. M. *J. Am. Chem. Soc.* **1989**, *111*, 6354.

(100) Goss, R. J. M.; Newill, P. L. A. *Chem. Commun. (Cambridge, U. K.)* **2006**, 4924.

(101) Winn, M.; Roy, A. D.; Grueschow, S.; Parameswaran, R. S.; Goss, R. J. M. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4508.

(102) Budisa, N.; Rubini, M.; Bae, J. H.; Weyher, E.; Wenger, W.; Golbik, R.; Huber, R.; Moroder, L. *Angewandte Chemie International Edition* **2002**, *41*, 4066.

(103) Miles, E. W.; Bauerle, R.; Ahmed, S. A. *Methods Enzymol.* **1987**, *142*, 398.

(104) Lanzetta, P. A.; Alvarez, L. J.; Reinach, P. S.; Candia, O. A. *Anal. Biochem.* **1979**, *100*, 95.

(105) Davisson, V. J.; Woodside, A. B.; Poulter, C. D. *Methods Enzymol.* **1985**, *110*, 130.

(106) Davisson, V. J.; Woodside, A. B.; Neal, T. R.; Stremler, K. E.; Muehlbacher, M.; Poulter, C. D. *J. Org. Chem.* **1986**, *51*, 4768.

(107) Rajagopalan, S.; Radke, G.; Evans, M.; Tomich, J. M. *Synth. Commun.* **1996**, *26*, 1431.

(108) Yamada, T.; Lukac, P. J.; Yu, T.; Weiss, R. G. *Chem. Mater.* **2007**, *19*, 4761.

(109) Gerstenberger, B. S.; Konopelski, J. P. *J. Org. Chem.* **2005**, *70*, 1467.

(110) Herzner, H.; Kunz, H. *Tetrahedron* **2007**, *63*, 6423.

(111) Hettwer, S.; Sterner, R. *Journal of Biological Chemistry* **2002**, *277*, 8194.

(112) Soderberg, B. C.; Shriver, J. A. *J. Org. Chem.* **1997**, *62*, 5838.

(113) Kozikowski, A. P.; Ishida, H.; Chen, Y.-Y. *J. Org. Chem.* **1980**, *45*, 3350.

(114) Van, d. B. E. M. M.; Jansen, F. J. H. M.; De, G. A. T. J. W.; Baldew, A. U.; Lugtenburg, J. *Recl. Trav. Chim. Pays-Bas* **1990**, *109*, 287.

(115) Ma, J.; Yin, W.; Zhou, H.; Liao, X.; Cook, J. M. *J. Org. Chem.* **2009**, *74*, 264.

(116) Roy, M.; Keblawi, S.; Dunn, M. F. *Biochemistry* **1988**, *27*, 6698.

(117) Koide, T.; Otaka, A.; Suzuki, H.; Fujii, N. *Synlett* **1991**, 345.

(118) Numata, T.; Awano, H.; Oae, S. *Tetrahedron Lett.* **1980**, *21*, 1235.

(119) Obata, T.; Yoshikawa, S.; Taniguchi, H.; Ando, T. *Pept. Chem.* **1996**, *33rd*, 489.

(120) Touaibia, M.; Desjardins, M.-A.; Provencal, A.; Audet, D.; Medard, C.; Morin, M.; Breau, L. *Synthesis* **2004**, 2283.

(121) Blaszczyk, A.; Elbing, M.; Mayor, M. *Org. Biomol. Chem.* **2004**, *2*, 2722.

(122) Kragulj, E. J.; Gustafson, J. L.; Grotjahn, D. B. *Synlett* **2007**, 2851.

(123) Martin, M. T.; Thomas, A. M.; York, D. G. *Tetrahedron Letters* **2002**, *43*, 2145.

(124) Uchiro, H.; Kobayashi, S. *Tetrahedron Lett.* **1999**, *40*, 3179.

(125) Zheng, N.; McWilliams, J. C.; Fleitz, F. J.; Armstrong, J. D., III; Volante, R. P. *J. Org. Chem.* **1998**, *63*, 9606.

(126) Clavier, S.; Rist, O.; Hansen, S.; Gerlach, L.-O.; Hoegberg, T.; Bergman, J. *Org. Biomol. Chem.* **2003**, *1*, 4248.

(127) Shieh, W. C.; Carlson, J. A. *J. Org. Chem.* **1992**, *57*, 379.

(128) Obrecht, W.; Plesch, P. H. *Makromol. Chem.* **1981**, *182*, 1459.

(129) Torisu, K.; Kobayashi, K.; Iwahashi, M.; Egashira, H.; Nakai, Y.; Okada, Y.; Nanbu, F.; Ohuchida, S.; Nakai, H.; Toda, M. *Eur. J. Med. Chem.* **2005**, *40*, 505.

(130) Sloan, M. J.; Phillips, R. S. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1053.

(131) Qian, Q.; Schultz, A. W.; Moore, B. S.; Tanner, M. E. *Biochemistry* **2012**, *51*, 7733.

(132) Yu, X.; Li, S.-M. *ChemBioChem* **2011**, *12*, 2280.

(133) Yu, X.; Xie, X.; Li, S.-M. *Applied Microbiology and Biotechnology* **2011**, 1.

(134) Kremer, A.; Li, S.-M. *Microbiology (Reading, U. K.)* **2010**, *156*, 278.

(135) Zou, H.-X.; Xie, X.; Zheng, X.-D.; Li, S.-M. *Applied Microbiology and Biotechnology* **2011**, *89*, 1443.

(136) Goloubinoff, P.; Gatenby, A. A.; Lorimer, G. H. *Nature (London)* **1989**, *337*, 44.

(137) Raghava, G. P. S. *CASP4* **2000**, 75.

(138) Myers, E. W.; Miller, W. *Comput. Appl. Biosci.* **1988**, *4*, 11.

(139) Eswar, N.; Webb, B.; Marti-Renom Marc, A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.-Y.; Pieper, U.; Sali, A. *Curr Protoc Bioinformatics* **2006**, *Chapter 5*, Unit 5 6.

(140) Tatusova, T. A.; Madden, T. L. *FEMS Microbiol. Lett.* **1999**, *174*, 247.

(141) Wenkert, E.; Sliwa, H. *Bioorg. Chem.* **1977**, *6*, 443.

(142) Majmudar, J. D.; Gibbs, R. A. *ChemBioChem* **2011**, *12*, 2723.

(143) Ruan, H.-L.; Yin, W.-B.; Wu, J.-Z.; Li, S.-M. *ChemBioChem* **2008**, *9*, 1044.

(144) Brown, H. C.; Okamoto, Y. *J. Am. Chem. Soc.* **1958**, *80*, 4979.

(145) Sangster, A. W.; Stuart, K. L. *Chem. Rev.* **1965**, *65*, 69.

(146) Wollinsky, B.; Ludwig, L.; Xie, X.; Li, S. M. *Org Biomol Chem* **2012**, *10*, 9262.

(147) Viswanathan, R., Personal communication.